THE FIFHT INTERNATIONAL WORKSHOP
CONTROL & INFORMATION TECHNOLOGY

Organised by
Institute of Electronics
Faculty of Automatic Control, Electronics and Computer Science
Silesian University of Technology

EAN-9788390840994

Gliwice, 21-22 September 2006
Programme Committee
Jan Bílek (Czech Republic)
Krzysztof Cetnarowicz (Poland)
Zdzisław Filus (Poland) - chairman
Vladimír Haasz (Czech Republic)
Edward Hrynkiwicz (Poland)
Ladislav Jurišica (Slovak Republic)
Peter Kopacek (Austria)
Štefan Kozák (Slovak Republic)
Edward Nawarecki (Poland)
Vladimír Řeřucha (Czech Republic)
Jerzy Rutkowski (Poland)
František Šolc (Czech Republic)
Vilém Srovnal (Czech Republic) - vice-chairman
Wiesław Wajs (Poland)
František Zezulka (Czech Republic)

Organising Committee
Prof. Zdzisław Filus
Łukasz Chruszczek
Marcin Kucharczyk
Mrs. Maria Drelichowska – secretary

Venue of the Workshop
Silesian University of Technology
Faculty of Automatic Control, Electronics and Komputer Science
Institute of Electronics
Ul. Akademicka 16, 44-100 Gliwice, Poland

The papers included in these proceedings were printed, in general, from camera-ready materials submitted by the authors. In some cases minor modifications have been made to adjust the papers to the required format.
FOREWORD

On behalf of the Organising and Programme Committees it is my great honour and pleasure to welcome the participants of the Fifth International Workshop Control & Information Technology IWCIT’06 to the Faculty of Automatic Control, Electronics and Computer Science, part of the Silesian University of Technology in Gliwice.

The first two and fourth International Workshops Control & Information Technology were organised by the Department of Measurement and Control, Technical University of Ostrava, Faculty of Electrical Engineering and Computer Science, Czech Republic, in 1999, 2001 and 2005. Thanks to close relationships between this department and the Institute of Electronics at the Silesian University of Technology the organisation of 2003’s and this year’s workshop was entrusted to our Institute.

Thirty eight papers have been accepted for this workshop. They will be presented in the following oral sessions:

- Control Systems
- Electronics
- Information Technology
- Telecommunication

We hope that this workshop, like the preceding ones, will prove once more to be useful forum for exchange of your ideas and experiences, helping you to complete your PhD theses.

We wish you all fruitful and inspiring discussions during the Workshop and pleasant stay at the Silesian University of Technology.

Zdzisław Filus
Contents

Section 1 – Control Systems

Traian CHIULAN, Brândușa PANTELIMON
On-line Monitoring of Large Power Transformer Units ........................................... 2

Paula DEACONESCU, Sanda Victorinne PATURCA
Scalar command using FEM model of the induction machine ................................ 8

Teresa GŁÓWKA
Mixed-order estimator .................................................................................................. 14

Pavel HOŘÍNEK
The computation of the electric characteristics in the overhead power transmission line with LabView ................................................................. 20

Mariusz LATOS
Active Noise Controller Based on Logarithmic Number System .............................. 27

Szymon OGONOWSKI
Control algorithms for heating system in small building ........................................ 33

Sanda Victorinne PATURCA, Paula DEACONESCU
Direct Torque Control with Reduced Torque Ripple for Permanent Magnet Synchro-nous Motor ................................................................................................................. 39

Michał PYŁA
Three-dimensional sound intensity field measurement with Matlab and Siglab 20-42 ... 45

Elena Ioana VĂTĂJELU
Identifying an electric oven with two heating zones .................................................. 51

Section 2 – Electronics

Łukasz CHRUSZCZYK, Damian GRZECHCA, Jerzy RUTKOWSKI
Fault diagnosis of analog electronic circuits ............................................................... 58

Weronika IZYDORCYZK
Computational approaches to the electrical properties of the tin dioxide surface – effect of oxygen adsorption ................................................................. 65

Weronika IZYDORCYZK
Response of SnO₂ - based gas sensor structure to synthetic air and water vapour ...... 70
Pavel KUBINA
*Laboratory thermo regulator for measuring the temperature dependence of biomedical sensors* ........................................................................................................................................................................ 76

Ireneusz KUROWSKI, Jerzy RUTKOWSKI
*Use of modular neural networks to fault diagnosis in analog active filters* .................. 80

Pavel NEVRIVA, Premysl PLESIVCAK
*Mathematical simulations of heat transfer in electrical sensors* ............................ 86

Marcin SZELEST
*Design of complex digital circuits in aspect of disclosure emission with use of simplified mathematic emanation model* ........................................................................................................................................................................ 90

**Section 3 - Information Technology**

Paweł BADURA
*Adaptive Thresholding in Fuzzy Approach to Segmentation of Cruciate Ligaments* ........ 97

Mariusz BAK
*The Application of Support Vector Machines and Kernel Methods in Speaker Recognition* ........................................................................................................................................................................ 103

Łukasz ĆMIEŁOWSKI, Patryk WALAWENDER
*Application of PCA algorithm for DNA microarrays gene expressions* ....................... 109

Rafał DONIEC
*The optimization of measuring method of blood glucose from dry bar code tests* ........ 115

Corneliu Octavian DUMITRU, Inge GAVAT
*Statistical and Hybrid Modeling for Digits Recognition in Romanian Language* ........... 119

Daniela FAUR, Mihai DATCU, Inge GAVAT
*Algorithms Evaluation for Remote Sensing Images Segmentation* ............................. 125

Daniela FAUR, Mihai DATCU, Inge GAVAT
*Information Theory Measures for Changes Detection in Earth Observation Images* .... 131

Aleksandra GRUCA
*Algorithms for inducing decision rules based on rough sets theory* ......................... 138

Przemysław KACPRZAK
*Joint trade of flowgate rights* ......................................................................................... 145

Katerina NOVAKOVA, Petr MARTINCIK, Miroslav POKORNY
*Fuzzy-wavelet based image compression* ....................................................................... 151

Pavel PETRÁNEK, Abdelbaki DJOUAMBI, Alina BESANÇON-VODA
*Using Genetic Algorithm for Fractional Systems Identification* .................................. 156
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application of fuzzy logic for gene expression analysis and classifications</td>
<td>162</td>
</tr>
<tr>
<td>3D Segmentation of Anatomical Structures Applied Fuzzy Classification</td>
<td>168</td>
</tr>
<tr>
<td>Comparative Analysis of Classical PI Speed Controller and Fuzzy Controller Used in Direct Torque Control of Induction Motor</td>
<td>174</td>
</tr>
<tr>
<td>Localization of the Posterior Cruciate Ligament on the T1-weighted MR Knee Images</td>
<td>180</td>
</tr>
<tr>
<td>Registration of the T1- and T2-weighted MR Knee Images</td>
<td>186</td>
</tr>
<tr>
<td>Use of data mining methods to improve sleep/wake stages classification</td>
<td>192</td>
</tr>
</tbody>
</table>

**Section 4 – Telecommunication**

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stochastic models for multiplexed VoIP traffic</td>
<td>199</td>
</tr>
<tr>
<td>Admission control for EDCA in 802.11 wireless networks</td>
<td>205</td>
</tr>
<tr>
<td>ZigBee in Biotelemetry</td>
<td>212</td>
</tr>
<tr>
<td>USB real-time data acquisition module for ultrasound scanner</td>
<td>217</td>
</tr>
<tr>
<td>Performance of Reduced Complexity Decoding of LDPC Codes with Few Quantization Bits</td>
<td>222</td>
</tr>
</tbody>
</table>
Section 1

Control Systems
On-line Monitoring of Large Power Transformer Units

Traian CHIULAN and Brândușa PANTELIMON
National Energy Research Institute – ICEMENERG, Bucharest, Romania
Polytechnic University, Bucharest Romania
ctraian@gmail.com

Abstract: The paper presents a software application representing a module of the application package utilized by a transformer unit monitoring system. The application has been developed by means of virtual instrumentation and enables recording of the transformer unit monitored parameters. The data are stored in files and can be subsequently analyzed and viewed or stored in an Excel compatible file. Thus, the data can be taken over by other applications such as the expert systems, or included into different other reports.

1 Introduction

The old age of the transformer substation primary equipment, correlated with the energy sector restructuring, understaffing and increase in competition on the services market, requires the introduction of new solutions and criteria for the correct evaluation of the momentary technical condition of the equipment in the transformer substations for its correct diagnosis, dimensioning of necessary maintenance work amount and their efficiency and the failure risk.

In order to increase availability and obtain an optimum management of the primary equipment management in the transformer substations it is imperiously necessary to introduce real time monitoring systems. These systems are absolutely necessary for the remote management of the electric substations.

The modern monitoring systems are based on distributed measurement system architecture, including physical process interfaces providing primary quantities, information control and digital processing blocks and the user interface.

The data taken over from the monitoring system should be transmitted to the control room of the electric substation and then, further to the territorial and national dispatcher, as well as to the organizations that assures their maintenance. An evaluation of the monitored equipment condition can be performed on the basis of the data provided by the monitoring systems by means of mathematical modeling, and by recording the data in a database the evolution of the equipment can be monitored in the course of time. These aspects are very useful for the maintenance activities carried out for these pieces of equipment.

2 Presentation of the monitoring equipment

The monitoring equipment has been manufactured by the US Qualitrol Corporation and is part of the series 509 – 100. This system has been already mounted in substations belonging to CN Transelectrica SA, such as: Slatina, Gutinas, Dumbrava, Lacul Sarat, Resita, Pestis, Stuparei, Pitești Sus, Brazi Vest, as well as electricity discharge substations belonging to CN Termoelectrica SA Company from the Turceni, Rovinari and Mintia CHP plants.
The front panel of the monitoring system is presented in Figure 1, and the interconnection diagram in Figure 2.

This system is an IED (Intelligent Electronic Device) mainly for monitoring the temperature of the transformer unit winding. When certain thresholds imposed by the user are surpassed the system signals this through the relay switching.

The system has 8 analogical inputs to which diverse sensors can be connected such as: thermal resistance of the PT100 type for the oil temperature measurement, current transformers for the measurement of the load current intensity, voltage transformers for the measurement of voltage both on the high voltage and the low voltage sides, sensors for the measurement of the oil level in the conservator, etc.

Input current blocks (4 – 20mA) that are compatible with EMS – SCADA equipment can also be mounted. Through these inputs information from other monitoring systems can be taken over such as those for the monitoring of the gas content in oil such as Hydran, or of the water content such as Waisalla.

Based on the information from the temperature and current sensors the equipment calculates the winding temperatures, as well as the lifetime consumption according to the thermal model. The user sets the parameters of the thermal model based on the data provided by the manufacturer. The user can set various hot spot thresholds. When the thresholds are reached, it switches one of the 8 relays the monitoring system is equipped with. This way the cooling batteries can start operating, certain warning signals can be given, or sometimes even the transformer unit can be disconnected from the network. Even a histeresis can be set, so that to prevent the repeated relay switching due to the small variations of the quantities around the threshold value.
The system is also equipped with a supplementary relay for the heating system control. Figure 2 illustrates the fact that the monitoring system has remote data transmission capabilities. For this, 4 unified signal (4 – 20mA) SCADA compatible outputs (RTU outputs) and a serial port of the RS – 485 type (current loop) for data transmission over long distances (up to 1.5 km) have been provided.

The RS-232 port on the front panel is provided only for the monitoring system configuration and not for remote data transmission. The system has been configured by means of the Remotes software provided by the manufacturer.

3 Presentation of the software

The equipment has been accompanied by an application that “brings” the front panel of the monitoring system on the user’s computer screen, as in Figure 3. By means of this application the user can “push” the buttons on the front panel of the monitoring system from a remote location. At the same time, what is displayed on this system monitor can be visualized. The main disadvantage of this application is the impossibility to record the evolution of the monitored parameters with time.

In order to eliminate this disadvantage, an application enabling recording (logging) of the monitored parameters in the course of time has been developed. The user interface is presented in figure 4. The application has been developed by means of the virtual instrumentation, utilizing the LabVIEW development environment.

The main window of the application presents the evolution of the 8 input quantities in the course of time, over a period of time of 30 minutes. The quantities are then “scrolled” on the left side of the screen, so that their evolution can be permanently seen during a half an hour time interval. The data acquired are updated every 15 seconds, this value having an inferior limit due to the communication speed of the serial port. For each of the 8 parameters 3 require / receive data sequences are necessary.
The application enables the export in the graphic format of the 8 diagrams. The export can be performed (made):

- to the clipboard memory, for their direct inclusion into various reports developed, for example in Microsoft Word;
- to the graphic files, in the following formats: bitmap (bmp), Encapsulated Postscript (eps) and Enhanced Metafile (emf).

Figure 5 illustrates such an example of export in the graphic format.

In this example only 3 parameters have been recorded, namely:

- the oil temperature (input 1);
- load current (input 2);
- oil level in the conservator (input 7).

The parameters of the communication port, the address of the monitoring equipment (as the RS – 485 port enables the connection of several pieces of equipment to the same communication path it is necessary that each equipment have its own address by means of
which it is identified), as well as the path where the data file will be stored are set by means of the *Options* window presented in Figure 6.

The *Options* window also enables the configuration of the logged parameters, as shown in figure 7. The following options can be set for each parameter:

- ✓ the name of the monitored parameter;
- ✓ whether it has to be stored in the file, or not;
- ✓ if the monitored parameter comes from a current transformer.

The data are recorded in a Microsoft Excel compatible file (Figure 8), enabling data taking over by expert systems, or by other EMS – SCADA software applications, as well as the development of diagrams with a view to introducing them into various reports. In this file both the current value of the parameter and its maximum and minimum values are recorded. This way the limits between which the parameter has varied can be also known. These limits can be reset by resetting the monitoring system by means of a button on its front panel.

In case the operation of the computer is temporarily interrupted out of various reasons, such as an electricity supply failure, interrupting the recording process, as well, the moment when the recording is resumed the data will be stored in the same file as before the interruption, exactly where it stopped.
4 Conclusions

This article presents a software application enabling the recording of the monitored parameter variation with time for a transformer unit. The application has been developed by means of virtual instrumentation, due to the latter’s flexibility as well as to the rapidness of the developed applications. The utilized development environment is LabVIEW.

The application makes it possible both to store the variation of the monitored parameters in graphic form, and the data in a Microsoft Excel compatible file. This facilitates the data taking over by other applications such as the expert systems, EMS – SCADA, or those for the report generations.

This application represents a module from a larger software package utilized by such a monitoring system.

The development of other modules, such as:
- module for the record management;
- module for the record analysis;
- module for the report automatic generation
are envisaged in the future.

5 References

Scalar command using FEM model of the induction machine

Paula Deaconescu, Sanda Victorinne Paturca
"Politehnica" University of Bucharest
paula_d@amotion.pub.ro, sanda@amotion.pub.ro

Abstract: This paper presents the finite element model of the induction machine based on the scalar command leading to the assessment of the dynamic performance of closed-loop ac drive system. The finite element model is characterized by the physical properties of the materials and of the electrical properties, taking into account the influence of the skin effect, stator - rotor slots combination, magnetization curve on the operating conditions of the machine. It is made a comparison of the finite element field model and the model with constant electrical parameters. The model with constant parameters is based on the transient equations with constant parameters of the induction machine and is implemented in Matlab / Simulink.

1 Scalar command scheme of the induction machine with constant parameters

The principle of the scalar control is to adjust the motor angular velocity $\omega$ by varying the stator frequency $f_s$ such as the magnetic state of the machine is about fixed, through the preservation of stator flux $\Psi_s$ in a constant value, and such as the torque follows the wished law of variation according to the speed. The simplicity of this type of control favors its implementation in many industrial devices conceived for these types of application.

The scalar command scheme of the induction machine with constant parameters is presented in the following figure.

Fig. 1 Scalar command scheme of the induction machine with constant parameters
The scheme contains:
- the reference value of the angular velocity 314 rad/s;
- a PI regulator for the angular velocity, which values are: \( K_p = 2 \), \( K_i = 1 \), saturation = \( \pm 3.5 \). It has as input the difference between the reference and feedback angular velocity and as output the rotor pulsation \( \omega_r \);
- a subsystem of the scalar command;
- a subsystem of the induction machine.

The induction machine subsystem has as inputs the three sinusoidal voltages \( u_a, u_b, u_c \) and as outputs the currents \( i_a, i_b, i_c \), the rotor flux components in the fixed coordinate system \( \Psi_{rds}, \Psi_{rfr} \), the magnetic torque and the angular velocity.

The subsystem of the scalar command is presented in the following figure.

The subsystem has as inputs the rotor and mechanical pulsations. The stator pulsation is obtained by addition of these ones:
\[
\omega_s = \omega_m + \omega_r.
\]

The stator flux \( \Psi_s \) is kept constant having the value
\[
|\Psi_s| = \frac{U_s - \text{peak}}{\omega_s} = \frac{536}{2 \times \pi \times 50} = 1.7 \text{ Wb}.
\]

The equivalent scheme parameters of the induction machine \( L_m, L_r, R_r, L_s, R_s \) are constant, independent of the operation conditions.

2 Scalar command scheme of the induction machine with variable parameters

In this section, the induction machine model is a finite element model. The scalar command scheme of the induction machine with variable parameters is presented in Fig. 3. The only difference in the scalar command scheme compared with the one from the previous section is the induction machine model. The finite element model takes into account the magnetization curve of the materials, the stator-rotor slots combination and the skin effect. The stator and rotor magnetic cores made of laminations are nonlinear with a saturation flux density of \( B_s = 2.1 \text{ T} \). The stator armature contains 24 slots and the rotor armature 20 slots.
IWCIT’06

Fig. 3. Scalar command scheme of the induction machine with variable parameters

Boundary conditions are of Dirichlet type on the two semi-circles delimiting the computation domain, i.e. outer surface of the stator magnetic core and inner surface of the rotor magnetic core in contact with the motor shaft, and Cyclic type on the two horizontal lines closing the computation domain, Fig. 4.

Fig. 4. Finite element field model of the induction machine

Due to the symmetry, the all computation domain consisting in the geometry of the induction machine with two poles, can be divided by 2, such as the geometry represents one pole and the corresponding computation time is decreased.

A complete finite element analysis of induction motor considers the coupling between the field model and the electrical circuit, Fig. 5.

The circuit model contains:
- three voltage sources, $U_{fV}$, $U_{fV}$, $U_{fw}$ characterized by 0 V, being controlled by Simulink;
- the leakage inductance of winding frontal sections outside the stator slots $L_{sf} = 2.02 \times 10^{-3}$ H;

Fig. 5. The associated electrical circuit
• coils of stranded conductor type, series connected, on each phase of the motor, U1, U2, V1, V4 and W2, W3. Each coil has the number of conductors \( N = \frac{w_s}{2} = 104 \) and the resistance \( 0.385 \Omega \). This last value is \( 1/4 \) of the phase resistance \( R_s = 1.54 \Omega \);
• a macro component modeling the rotor electric circuit of squirrel cage type.

3 Results

The results are presented comparatively for the two models of the induction machine.

From the time variation of the phase currents it is observed that the rms value of the stationary regime is 2.7 A in the model with constant parameters and 8.2 A in the model with variable parameters. The stationary rms value of the stator currents in the model with variable parameters is 3 times greater than in the model with constant parameters. In Fig. 7 are presented the results of the mechanical angular velocity PI regulator for the first and second models of the induction machine.
It can be noticed the reference angular velocity at 314 rad/s, its feedback and the command of the PI regulator. Also, it can be observed that the regulator works well, because when the speed reaches the reference value the command of the regulator changes. In the following figures are presented the time variation of the electromagnetic torque for both models.

![Fig. 8. Time variation of the electromagnetic torque for the model with constant, (a), and variable, (b), parameters](image)

It can be observed that in the model with variable parameters, the electromagnetic torque has ripples due to the $L_m$ and $R_r$ parameters variations. Also, the electromagnetic torque in the induction machine model with variable parameters is with 50 % greater than in the model with constant parameters.

The time variations of phase voltages are presented in Fig. 9.

![Fig. 9. Time variation of the stator voltages for the model with constant, (a), and variable, (b), parameters](image)

The rms value of the phase voltage is 380 V. The transient regime is with approximately 0.5 s longer in the model with constant parameters than in the model with variable parameters.

The asking of the scalar command $\frac{U}{f} = ct$ is accomplished. When the voltage reaches the rated rms value 380 V, the frequency reaches the rated value 50 Hz and it rests constant.
In Fig. 10 is presented the time variation of the stator frequency $f_s$.

![Graphs showing time variation of stator frequency](image)

Fig. 10. Time variation of the stator frequency for the model with constant, (a), and variable, (b), parameters

### 4 References


Mixed-order estimator

Teresa Główka
Institute of Automatic Control, Silesian University of Technology
ul. Akademicka 16, PL-44-100 Gliwice, Poland
contact phone: +48 32 237 23 98, fax: +48 32 237 21 27
e-mail:teresa.glowka@polsl.pl

Abstract: In the paper a new identification method, employing second- and higher-order spectra, is considered. Proposed heuristic estimator, so-called mixed-order estimator, joins the advantages of both second- and higher-order identification methods. In this approach the frequency response estimator is a linear combination of classical and higher-order estimators, with weights dependent on the coherence function. The aim of the paper is to examine properties of the mixed-order estimator. Obtained estimates are computed on the basis of data derived from different simulation and laboratory experiments.

1 Introduction

Nowadays a nonparametric identification based on higher-order spectra (HOS) estimation is often and often exploited instead of classical spectral analysis based on estimation of second-order spectra (SOS), i.e. power spectral density. HOS contain some additional information about processed signals not available to extract using SOS. The most important property of HOS follows from their definitions: they are identically zero for Gaussian signals. Therefore in some cases HOS-based methods let us theoretically eliminate, in practice substantially reduce, an influence of additive Gaussian noise on identification results [1], [7], [8]. This is true also for non-Gaussian disturbances if we employ signal averaging described in [4]. On the other hand, the variance of the classical second-order estimates, outside the dominant frequency band of disturbance, is usually smaller than the variance of the higher-order ones, see Fig 1. In the paper a new heuristic estimator, so called mixed-order estimator, is proposed, that joins the advantages of both second- and higher-order identification methods. In the proposed approach a linear combination of classical and higher-order estimators is used, with weights dependent on the power of disturbances in the function of frequency. In the light of this statement it is obvious that the weights should exploit coherence function.

Fig. 1 Second- and higher-order frequency response estimates of the exemplary system with disturbances concentrated in the frequency band from 0.5 to 1.5.
2 Mixed-order estimator

The idea of proposed mixed-order estimator is as follows: let’s take into account results obtained using both higher-order and classical identification methods, and make an estimator being a linear combination of them. Let’s weight the both components of combination with a coefficient dependent on frequency, to assure that “more of the higher-order estimate” is taken for a frequency bands where the signals are strongly disturbed, and “more of the second-order estimate” outside these bands. Simplifying the above heuristic rule, we can say that for the system illustrated in Fig. 1 the higher-order estimate is more meaningful for frequency band 0.5 to 1.5, and second-order estimate elsewhere. The mixed-order estimator and both its components are described below.

2.1 Second-order estimator

The classical spectral analysis is the most known nonparametric identification method. Given system input $u(i)$ and output $y(i)$ data sequences ($i = 0, 1, \ldots, N-1$), the frequency response estimator $\hat{K}_{SO}(j\Omega n)$ is expressed with help of second-order spectra as

$$
\hat{K}_{SO}(j\Omega n) = \left( \frac{\hat{\Phi}_{uu}(j\Omega n)}{\hat{\Phi}_{uu}(j\Omega n)} \right)^* ,
$$

for frequencies $\Omega n$ ($\Omega = 2\pi/N, n = 0, 1, \ldots, N-1$) [6], [9]. The estimates of input power spectral density $\hat{\Phi}_{uu}(j\Omega n)$ and input-output cross-power spectral density $\hat{\Phi}_{uy}(j\Omega n)$ in the direct approach (periodogram) can be calculated as a multiplication of the proper $N$-point discrete Fourier transforms of $u(i)$ and $y(i)$

$$
U(j\Omega n) = \sum_{i=0}^{N-1} u(i)e^{-j\Omega n i} , \quad Y(j\Omega n) = \sum_{i=0}^{N-1} y(i)e^{-j\Omega n i} ,
$$

according to the equations

$$
\hat{\Phi}_{uu}(j\Omega n) = \frac{1}{N} U(j\Omega n)U^*(j\Omega n) , \quad \hat{\Phi}_{uy}(j\Omega n) = \frac{1}{N} U(j\Omega n)Y^*(j\Omega n) ,
$$

where superscript * denotes the complex conjugation.

2.2 Higher-order estimator

The higher-order identification method used in the paper is based on direct estimators of integrated bispectra (IB) [8], [10]. The method was chosen for its good noise-reduction properties and small computational complexity. In the IB method, the frequency response estimator $\hat{K}_{HO}(j\Omega n)$ is a ratio of integrated cross-bispectrum of input and output signals to the integrated bispectrum of input signal, according to the equation

$$
\hat{K}_{HO}(j\Omega n) = \left( \frac{\hat{B}_{uy}(j\Omega n)}{\hat{B}_{uu}(j\Omega n)} \right)^* ,
$$

The direct estimates of IB are calculated on the basis of discrete Fourier transforms of signals $u(i), y(i)$ and $r_{2u}(i)$ as

$$
\hat{B}_{uu}(j\Omega n) = \frac{1}{N} R_{2u}(j\Omega n)U^*(j\Omega n) , \quad \hat{B}_{uy}(j\Omega n) = \frac{1}{N} R_{2u}(j\Omega n)Y^*(j\Omega n) ,
$$
IWCIT’06

where

\[
U(j\Omega) = \sum_{i=0}^{N-1} u(i)e^{-j\Omega i}, \quad Y(j\Omega) = \sum_{i=0}^{N-1} y(i)e^{-j\Omega i}, \quad R_{2u}(j\Omega) = \sum_{i=0}^{N-1} r_{2u}(i)e^{-j\Omega i}.
\] (6)

Signal \( r_{2u}(i) \) is defined as

\[
r_{2u}(i) = u^2(i) - \frac{1}{N} \sum_{k=1}^{N} u^2(k).
\] (7)

Unlike the classical second-order identification methods, the IB-based method allows to obtain strongly consistent estimates in the presence of disturbance [10]. The IB estimates obtained in the direct way should be smoothed in the frequency domain, otherwise we receive the empirical transfer function estimator [6], and the method order decreases to the second.

2.3 Mixed-order estimator

In the proposed approach, the frequency response estimator \( \hat{K}_{MO}(j\Omega) \) is a linear combination of classical and higher-order estimators, weighted by the \( \alpha(\Omega n) \) and \([1 - \alpha(\Omega n)]\) coefficient, respectively

\[
\hat{K}_{MO}(j\Omega) = \alpha(\Omega n)\hat{K}_{HO}(j\Omega) + [1 - \alpha(\Omega n)]\hat{K}_{HO}(j\Omega),
\] (8)

where \( \alpha(\Omega n) \) depends on the coherence function of \( u(i) \) and \( y(i) \)

\[
\kappa^2(\Omega n) = \frac{S_{yy}(j\Omega n)}{S_{uu}(\Omega n)S_{yy}(\Omega n)}.
\] (9)

Because of the heuristic nature of the proposed approach, there is no rule determining in advance how to choose the \( \alpha(\Omega n) \) coefficient. Some propositions can be suggested:

- Coherence
  \[
  \alpha(\cdot) = \kappa^2(\cdot)
  \] (10)

- Root of coherence
  \[
  \alpha(\cdot) = \sqrt{\kappa^2(\cdot)}
  \] (11)

- Scaled coherence
  \[
  \alpha(\cdot) = \frac{\kappa^2(\cdot) - \min\{\kappa^2(\cdot)\}}{1 - \min\{\kappa^2(\cdot)\}}
  \] (12)

that for the lowest value of coherence weight parameter becomes zero \( \alpha(\arg\min\{\kappa^2(\cdot)\}) = 0 \).

- The two-level coefficient which value is conditional on a sign between the coherence and a certain level \( p \)
  \[
  \kappa^2(\cdot) < p \Rightarrow \alpha(\cdot) = 0,
  \]
  \[
  \kappa^2(\cdot) \geq p \Rightarrow \alpha(\cdot) = 1
  \] (13)

Fig. 2 Illustration of differences between \( \alpha(\cdot) \) chosen according to the proposed rules for exemplary coherence function. \( \alpha(\cdot) \) is coherence (1), root of coherence (2), scaled coherence (3), two-level coefficient, \( p = 0.8 \) (4).
One should notice that in (8) the second- and higher-order components have to be determined for the same set of frequencies. Of course another higher-order method than described by (4)-(7) can be used. The frequency response estimate can be easily transformed to parametric model using e.g. the least squares approximation of frequency response for fixed model structure described in [10].

3 Identification results

In this section results of identification with the proposed approach are derived and compared with results obtained using second- and higher-order methods. The data were acquired in the simulation and laboratory experiments.

3.1 Simulation experiments

The higher-order identification methods are useful in two cases: (1) when the input measurement is noisy, and (2) when the same disturbance is correlated with the system input and output. Second situation occurs in all feedback and feedforward systems (e.g. in active noise control). During the experiment the system $K(z^{-1})$

$$K(z^{-1}) = \frac{0.8z^{-1} + 0.2z^{-2} - 0.1z^{-3}}{1 - 0.2z^{-1} + 0.02z^{-2} - 0.08z^{-3} + 0.2z^{-4} + 0.2z^{-5}}.$$  (14)

was simulated. Results of identification are illustrated in the form of root mean square error (RMSE) of frequency response estimates obtained in 100 Monte Carlo runs, see Fig. 3.

Case 1: noisy input and output, disturbances in the frequency range 0.5 to 1

Case 2: feedback system, disturbances in the frequency range 1.2 to 1.7

Case 3: feedforward system, disturbances in the whole frequency range with max for 1.25

Fig. 3 RMSE of second-order estimates (dotted line), higher-order estimates (dashed line) and mixed-order estimates with different $\alpha(\cdot)$ coefficient (grey lines). Results of nonparametric (left) and parametric (right) identification.
In the identification procedure firstly the frequency response estimates were obtained, and secondly the parametric models were computed and transformed again to the frequency domain. This is referred as nonparametric and parametric results, respectively. For the nonparametric results we can observe that mixed-order method works properly, producing error equal to error of higher-order method in the dominant frequency bands of disturbances, and equal to error of second-order method outside these bands. The mixed-order estimates are better if the disturbances are concentrated in some frequency bands and worse if the disturbances are scattered in the whole frequency band. For the parametric results it is not so easy to draw any general conclusion because the proposed method behaves once better, once worse, than the higher-order one. But we should notice that transformation of nonparametric into parametric model can be treated as a kind of averaging, therefore for models of low order we loose the advantages of mixed-order method in comparison to higher-order (lower variance outside the dominant frequency band). The same loss of advantages appears if we average models obtained in several experiments. The situation changes for identification of models of very high order, like in active noise control system.

3.2 Laboratory experiments

An example of on-line secondary path model identification in a feedforward active noise control (ANC) system, placed at the Institute of Automatic Control, Silesian University of Technology, Gliwice, Poland, is shown in the Fig. 4. The adaptive ANC system of interest, described in details in [2], [3], [4] creates a local spatial zone of quiet surrounding an error microphone in an enclosure. During the experiment a disturbance (attenuated noise) was concentrated in the frequency band from 70 to 110 [Hz]. Pattern of the secondary path was obtained in the additional off-line experiment in the undisturbed system.

During the identification procedure a parametric model (FIR filter) with one hundred coefficients was found. Identification of very high order models is caused by very complicated dynamic of the ANC system paths. Identification results in Fig. 4 show again the advantages of the mixed-order estimator in comparison to the second- and higher-order estimators.
4 Conclusion

In the paper a new heuristic mixed-order estimator, that joins the advantages of both second- and higher-order identification methods, was presented. In the proposed approach the frequency response estimator is a linear combination of classical and higher-order estimators, with weights dependent on the coherence function. The properties of the mixed-order estimator were shown with identification results. Obtained estimates were computed on the basis of data derived from different simulation and laboratory experiments. Using mixed-order estimator in some cases we gain reduction of identification error.

5 Acknowledgments

The partial financial support of this research by The Polish Ministry of National Education (grant 3 T11A 012 29) is gratefully acknowledged.

6 References

The computation of the electric characteristics in the overhead power transmission line with LabView

Pavel Hořinek
VŠB - Technical University of Ostrava
pavel.horinek@vsb.cz

Abstract: The paper describes one of the most possible utilization of the development environment LabView. The application created in LabView can be used to computation the electrical inductance L and capacitance C of the single and multiple outside lines.

1 The computation of the inductance and capacitance of the outside lines

1.1 The inductance of the lines

The inductance of each conductor at the line depends on its position on the tower. The transposition of the line or the line twist causes decreasing this dependence. The relationship of the inductance for the single transposition- and the twin(multiple) compensated transposition line is:

\[ L_k = 0,2 \cdot \ln \frac{a_s}{r} + 0,05 \cdot \mu_r \text{ (mH. km}^{-1}) \] (1)

and for twin(multiple) according transposition lines:

\[ L_k = 0,2 \cdot \ln \frac{a_s \cdot c_s}{r \cdot b_s} + 0,05 \cdot \mu_r \text{ (mH. km}^{-1}) \] (2)

where

- \( L_k \) (mH.km\(^{-1}\)) is the inductance one’s phase of an line,
- \( a_s \) (m) - conductors mean geometric distance from among diverse phase of the single line,
- \( b_s \) (m) - conductors mean geometric distance same phase of the diverse lines,
- \( c_s \) (m) - conductors mean geometric distance diverse phase of the diverse lines,
- \( r \) (m) – radius of conductor (Fig. 1),

Computation the radius of the conductor \( r \), which is in the previous relations is possible at the circle wire from the cross section, at the Aldrey Fe is need this value find in the tables [1], [2].

Fig.1 Distance of the lines on the transmission tower
1.1.1 The inductance of the bundle conductors

Bundle conductors are usually used at the lines of very and extra high voltage. This meaning that every phase is not created only by one line, but with bundle some conductors, which have from each other the distance \( a \). Number of the conductors in the bundle \( n \) is different.

In the light of electromagnetic field is the bundle conductor in principle as round conductor with the equivalent radius \( r_e \). For this radius holds next relationship:

\[
r_e = \sqrt[2]{r \cdot a_1 \cdot a_2 \cdots a_n} \quad (m)
\]  

(3)

For the mean inductance of the single line with the bundle conductors holds then next relation:

\[
L_k = 0.2 \cdot \ln \frac{a_s}{r_e} + \frac{0.05 \cdot \mu_s}{n} \quad (mH.km^{-1})
\]  

(4)

in case of double line (see relationship (4)) we must consider in addition mean geometric distance \( b_s \) and \( c_s \) according to relationship (2).

1.2 The capacitance of the lines

For working capacitance of the one phase of the line with transposition and without earth-wires holds approximate relation:

\[
C_{pk} = \frac{55.67}{\ln \frac{a_s}{r}} \quad (nF. km^{-1})
\]  

(5)

In the generated application we calculate with the aid of potential coefficients, that give more accurately result. The basic relations for the computation the capacitance of the transposed lines are in the Tab.1. The value calculates according to the next relationship:

\[
C_{pk} = \frac{10^{12}}{\Delta - \Delta'} \quad (nF. km^{-1})
\]  

(6)

Tab.1 The relations for capacitance of an outside lines with the aid of potential coefficient:

<table>
<thead>
<tr>
<th>Line</th>
<th>Without earth-wires</th>
<th>With earth-wires</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single</td>
<td>( \Delta = \delta )</td>
<td>( \Delta = \delta - \delta_z )</td>
</tr>
<tr>
<td></td>
<td>( \Delta' = \delta' )</td>
<td>( \Delta' = \delta' - \delta_z )</td>
</tr>
<tr>
<td>Double</td>
<td>( \Delta = \delta + \delta_x )</td>
<td>( \Delta = \delta + \delta_x - 2\delta_z )</td>
</tr>
<tr>
<td></td>
<td>( \Delta' = \delta' + \delta_x )</td>
<td>( \Delta' = \delta' + \delta_x - 2\delta_z )</td>
</tr>
</tbody>
</table>

where potential coefficients \( \delta \) (m.F^{-1}) represent band from among:

\( \delta \) – working conductors against earth,
\( \delta' \) – conductors of an single line,
\( \delta_{pq} \) – earth-conductors,
\( \delta_s \) – conductors same phase of an diverse lines,
\( \delta_x' \) – conductors diverse phase of an diverse lines,
\( \delta_w \) – working and earth-conductors,
\( \delta_{\text{ez}} \) – earth-conductors against earth,
\( \delta_z \) – represents total effect of the earth-conductors.

More information is possible find in the literature [1].

### 1.2.1 The capacitance of the lines with bundle conductors

For capacitance of the line with bundle conductors holds the same relations, which was presented, but in place of the radius of the conductor \( r \) gives value of the equivalent radius \( r_e \).

### 1.3 The conductor sag

For the computation of the inductance and capacitance of the outside lines we need to know maximal sag of the line, which is possible specify from equation or with smaller exigencies of the accuracy from equation of the parable. With accuracy sufficient for practice we can this compensation presume in the span 400 m. With larger span we must compute always with the span.

![Diagram of conductor sag](image)

**Fig.2** Primary data for computation of the sag

The maximal sag is in the centre of span. If we consider the equation of parable holds:

\[
f_{\text{max.}} = \frac{a^2 \gamma z}{8 \sigma_H} \quad (\text{m})
\]  

where

- \( a \) (m) is span
- \( \sigma_H \) (MPa) horizontal component tension of a conductor
- \( \gamma \) (N.m\(^{-1}\).mm\(^{-2}\)) specific weight of a conductor without icing
- \( z \) (-) overfreight owing to icing if you like wind

After computation the maximal sag we abate his two-thirds from high \( h \) of conductor on the transmission tower. We obtain the high of gravity centre of the conductor. This parametr we consider in computaton the inductance and capacitance of the line.
\[ h_{\text{typ}} = h - \frac{2}{3} f_{\text{max}} \text{ (m)} \]  

where \( h \) is high point of suspension conductor above the ground.

## 2 Development environment LabView

LabVIEW (short for Laboratory Virtual Instrumentation Engineering Workbench) is a platform and development environment for a visual programming language from National Instruments. The graphical language is named "G". Originally released for the Apple Macintosh in 1986, LabVIEW is used for data acquisition, instrument control, and industrial automation on a variety of platforms including Microsoft Windows, various flavors of UNIX, Linux, and Mac OS. The latest version of LabVIEW is version 8.0.

## 3 Description of application for computation of the electric characteristics in the overhead power transmission line

The solution algorithm of the created application is shown in flow diagram (see Fig. 3) Input data are scanned by interactive way from display (Front Panel).

![Diagram of solution algorithm](image)

Fig.3  Diagram of solution algorithm

After initiation the application displays front panel (Fig. 5). On the left side user entry with popup menu type of the conductor (Fig. 4). The created application includes current types, but it is possible to add next types of conductors. After the type of conductor entry and program start are generated on the left side of panel his characteristics,
namely:

d (mm)  conductor diameter,
m (kg/km) nominal conductor weight,
\(\sigma_H\) (MPa)  horizontal conductor tension,
S (mm²)  conductor cross-section,

On the left side of the panel there are beside three digital controls for span mast entry \(a\) and with icing on the line beside icing thickness \(t\) and icing density \(\gamma_n\).

On the right side panel selects user with menu mast type.

After type mast selection displays in the centre of panel the scheme with dimensions for

next technical characteristics entry:

\[h_1, h_2, h_3\text{ (m)}\] - conductors highs on the mast,
\[d_1, d_2, d_3\text{ (m)}\] - conductors distances from axe,
\(l\)  (km) - total line length.

Next input is the information, if the line of one phase is from single conductor or from bundle conductors. User do this with switch, which is situated under window with display of mast.

In case of selection bundle conductors, the next data have to be entered:

\[n\ (-)\] - conductors number in the bundle
\[a\ (cm)\] - conductors distance in the bundle

After all necessary input characteristics user turn on the application by the button “COMPUTATION”. After start of the program performs, according to above relations, computation of the electric characteristics L and C and the result displays in bottom of application panel. After press button “KONEC” erase all characteristics, inclusive of computed results and application waits on an new entry. The next two figures show the front panel and the block diagram of computaton.
Fig. 5  Front Panel

Fig. 6  One of the seven program source code for computation the electrical inductance and capacitance of the double lines with mast “fir”
4 Acknowledgement

The support for this research work has been provided by the GAČR project 102/05/H525: The Postgraduate Study Rationalization at Faculty of Electrical Engineering and Computer science VSB-TU Ostrava.
Active Noise Controller Based on Logarithmic Number System

Mariusz Latos
Silesian University of Technology
Mariusz.Latos@polsl.pl

Abstract: The paper presents some aspects of using logarithmic number system (LNS) in active noise control oriented applications. On the beginning the logarithmic number system is introduced and compared with alternative floating point number system (FP). Advantageous and weaknesses of both number systems are presented. Problems with changing number formats and implementation of basic arithmetical operations are discussed. Then Active Noise Controller (ANC) is created using LNS. In above example usefulness of LNS is tested and compared with FP arithmetic.

1 Introduction

In digital signal processing aspects computations are the most challenging. These computations often works in realtime, so large dynamic range of numbers used must be guaranteed. Two above restrictions are provided when floating point number system (FP) or logarithmic number system (LNS) are used. LNS has some advantageous in specific situations over the FP arithmetic. The principal advantage is easiness to implement and faster work multiplication and division operations. Of course there isn’t for free, it comes at the cost of complicated addition and subtraction operations. Using these feature, LNS arithmetic is applied in ANC controller. In the paper properties of ANC controller and influence of LNS on it are discussed and illustrated by simulations experiments.

2 Number systems comparison

2.1 Floating point number system

Floating point number representation ([9], [4], [7]):

\[ X = (-1)^{S_{FP}} \times 1.F_{FP} \times 2^{E_{FP}} \]  

where: \( S_{FP} \) – number’s sign, \( F_{FP} \) – number’s fraction, \( E_{FP} \) – number’s exponent

Floating point number format:

<table>
<thead>
<tr>
<th>Sign</th>
<th>Exponent</th>
<th>Significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 bit</td>
<td>e bits</td>
<td>f bits</td>
</tr>
</tbody>
</table>

2.2 Logarithmic number system

Logarithmic number representation ([7], [3], [2], [8]):

\[ X = (-1)^{S_{LNS}} \times 2^{E_{LNS}} \]  

<table>
<thead>
<tr>
<th>Sign</th>
<th>Exponent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 bit</td>
<td>e bits</td>
</tr>
</tbody>
</table>
where: $S_{LNS}$ – number’s sign, $E_{LNS}$ – fixed point number

Logarithmic number format:

Table 2 LNS number format

<table>
<thead>
<tr>
<th>flags</th>
<th>$S_{LNS}$</th>
<th>$E_{LNS}$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2 bits</td>
<td>sign</td>
<td>integer</td>
<td>fraction</td>
</tr>
<tr>
<td>1 bit</td>
<td></td>
<td>$e_1$ bits</td>
<td>$e_2$ bits</td>
</tr>
</tbody>
</table>

Flags are used to represent a special values, i.e. zero, $+\infty$, $-\infty$, NaN. Range and precision in LNS are similar to range and precision in FP ([3], [2]). Special values in FP numbers ([7]).

### 3 LNS implementation

#### 3.1 FP to LNS conversion

First step is checking if FP number is a special value and if yes, converting them into a equivalent in LNS ([7], [3]). Checking can be done by interpreting state of mantissa and exponent of FP number. Second step: the absolute value of created LNS number is ([7], [3]):

$$\log_2 (1.F_{FP} \times 2^{E_{FP}})$$

Of course sign of the LNS number is the same as FP number. (3) can be calculated as follows:

$$\log_2 (1.F_{FP} \times 2^{E_{FP}}) = \log_2 (1.F_{FP}) + \log_2 (2^{E_{FP}}) = \log_2 (1.F_{FP}) + E_{FP}$$

(4)

It is important that $E_{FP}$ is an integer and it is stored in memory in modified form. True integer portion of LNS number is $E_{FP}' = E_{FP} - \text{bias}$, so correction is needed (see [7]). Because of that there is $E_{FP}'$ instead of $E_{FP}$ in equation (4). It is noticeable that argument of nonlinear function (part of result’s sum of (4)) is in the range of [1,2) and function’s values are in the range of [0,1). In that range this function must be tabulated and stored in memory. Of course there is only subset of values stored in memory. So, to determine that function interpolation algorithm is needed. Conversion’s precision depends on quality of interpolation algorithm.

#### 3.2 LNS to FP conversion

First step is similar like previously ([7], [3]). In the second step the absolute value of created FP number is calculated. Sign of the FP number is the same as LNS number. Integer part of LNS number is directly the exponent of FP number. Of course before storing in the computer memory, bias modification is needed. Fraction part of LNS is in the range [0,1). It is the argument of the function:

$$2^N$$

(5)

where N is fraction part of LNS number. Values of the function (5) are in the range of [1,2), and this is directly the mantissa of FP number [7]. Function (5) must be tabulated and stored in memory in the selected range (similarly to function in previous section).

#### 3.3 Basic LNS arithmetic operations

Before every arithmetic operation flags of the numbers must be checked. If result becomes a special number then result’s flag must be set.
3.3.1 Multiplication and Division

Multiplication (Division) of two numbers in LNS ([7], [3], [2], [8]) is simply an addition (subtraction) LNS forms of 2 numbers (6), (8). Sign of the result is logical XOR product on the signs of 2 numbers (7).

\[
\log_2(C) = \log_2(A \times B) = \log_2(A) + \log_2(B) \tag{6}
\]

\[
sign(C) = sign(A) \text{ XOR } sign(B) \tag{7}
\]

\[
\log_2(C) = \log_2\left(\frac{A}{B}\right) = \log_2(A) - \log_2(B) \tag{8}
\]

3.3.2 Addition and Subtraction

Addition and subtraction are more complicated operations in LNS in opposition to multiplication and division ([7], [5], [3], [2], [8]). Assume:

\[
A = (-1)^{S_A} \times 2^{E_A}, \quad B = (-1)^{S_B} \times 2^{E_B}, \quad C = (-1)^{S_C} \times 2^{E_C} \tag{9}
\]

Assume relation between numbers A and B: |A| ≥ |B|. Then sign of the result is sign of number A (SC = SA). Absolute result’s value is:

\[
E_c = \log_2\left(|A \pm B|\right) = \log_2\left(A \left|1 \pm \frac{B}{A}\right|\right) = E_A + \log_2\left(1 \pm 2^{E_A-E_B}\right) = E_A + f(E_B - E_A) \tag{10}
\]

Summarizing, addition or subtraction of two numbers in LNS is sum of 2 values (10). First value is LNS form of number A and second, value of function \(f(x)\), for \(x = E_B - E_A\) (Fig. 1) ([7], [8], [2], [5]).

Function \(f(x)\) must be tabulated and stored in memory (similarly to stored functions in sections Sec.3.1 and Sec.3.2). In all examples and simulations in this paper, there is a linear interpolation used between 100 tabulated values. It is important, that for \(x = 0\) function \(f(x)\) for subtraction gains \(\pm\infty\) and this situation must be detected.

![Fig 1. Function f(x) for addition and subtraction](image)

4 Simulation of ANC system

In all simulations FP arithmetic will be used in single precision format and LNS arithmetic in equivalent of single precision (\(e_1 = 8\) and \(e_2 = 23\), Table 2).

4.1 ANC controller

There are many practical ANC applications([1], [6], [10], [11]). In this paper LNS arithmetic is used in implementation of ANC controller for creation of 3-D zone of quiet in enclosure.
There is feed forward structure model of ANC (Fig. 2). Frequency responses of selected ANC path models are presented in [10].

\[
J = -10 \log_{10} \left( \frac{\sigma_e^2}{\sigma_d^2} \right)
\]

where:
- \( x \) – noise
- \( Pe \) – disturbance path
- \( Px \) – measurement path
- \( S2 \) – secondary path
- \( W \) – finite impulse response (FIR) adaptive filter
- \( S1 \) – estimate of \( S2 \)
- \( LMS \) – adaptation algorithm

Fig. 2. ANC plant

FIR filter \( W \) together with adaptation algorithm \( LMS \) and \( S2 \) create a controller. Idea of the circuit is simple: controller produce a phase inverted signal in relation to \( d \). Sum of both signals produce the error signal \( e \). Adaptive control is necessary, because characteristic of the noise may change in time. Order of the filter \( W \) in all simulations was 400. \( S2 \) path introduce a time delay not taken into account in the LMS algorithm equations. \( S1 \) path compensate influence of \( S2 \). This resulting algorithm is called filtered \( x \) LMS (FX-LMS) [10]. The aim of the simulations is to show, how the influence of LNS arithmetic on the work of controller is. There will be computed level of attenuation in each case and minimal accuracy of numbers in LNS, which guarantee correct work of ANC controller. Decreasing of accuracy of LNS numbers is realized through cutting \( m \) less significant bits from fraction part of fixed point \( E_{LNS} \). The disturbance attenuation is calculated as [10]:

<table>
<thead>
<tr>
<th>Simulation number</th>
<th>Arithmetic</th>
<th>Tabulated values</th>
<th>Inactive bits</th>
<th>( \mu )</th>
<th>( \sigma_e^2 )</th>
<th>( \sigma_d^2 )</th>
<th>( J )</th>
<th>Figure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FP</td>
<td>( x )</td>
<td>( x )</td>
<td>0.001</td>
<td>2.267e-07</td>
<td>45.12</td>
<td></td>
<td>Fig. 3</td>
</tr>
<tr>
<td>2</td>
<td>FP</td>
<td>( x )</td>
<td>( x )</td>
<td>0.010</td>
<td>1.831e-07</td>
<td>46.05</td>
<td></td>
<td>Fig. 4</td>
</tr>
<tr>
<td>3</td>
<td>LNS</td>
<td>no</td>
<td>0</td>
<td>0.010</td>
<td>1.986e-07</td>
<td>45.70</td>
<td></td>
<td>Fig. 4</td>
</tr>
<tr>
<td>4</td>
<td>LNS</td>
<td>yes</td>
<td>0</td>
<td>0.010</td>
<td>2.388e-07</td>
<td>44.90</td>
<td></td>
<td>Fig. 4</td>
</tr>
<tr>
<td>5</td>
<td>LNS</td>
<td>yes</td>
<td>5</td>
<td>0.010</td>
<td>2.555e-07</td>
<td>44.61</td>
<td></td>
<td>Fig. 4</td>
</tr>
<tr>
<td>6</td>
<td>LNS</td>
<td>no</td>
<td>5</td>
<td>0.010</td>
<td>1.844e-07</td>
<td>46.02</td>
<td></td>
<td>Fig. 4</td>
</tr>
<tr>
<td>7</td>
<td>LNS</td>
<td>yes</td>
<td>5</td>
<td>0.001</td>
<td>1.731e-04</td>
<td>16.30</td>
<td></td>
<td>Fig. 5</td>
</tr>
<tr>
<td>8</td>
<td>LNS</td>
<td>no</td>
<td>5</td>
<td>0.001</td>
<td>2.260e-07</td>
<td>45.14</td>
<td></td>
<td>Fig. 3</td>
</tr>
<tr>
<td>9</td>
<td>LNS</td>
<td>yes</td>
<td>15</td>
<td>0.010</td>
<td>3.563e-06</td>
<td>33.16</td>
<td></td>
<td>Fig. 6</td>
</tr>
<tr>
<td>10</td>
<td>LNS</td>
<td>no</td>
<td>16</td>
<td>0.010</td>
<td>0.384e-00</td>
<td>-17.17</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.2 Simulations
Results from FP arithmetic are the base results, which are compared with results from LNS arithmetic. LMS parameter $\mu$ is increased from 0.001 to 0.01, because of slow disturbance attenuation for presented number of simulation samples (see Figures 3 and 4). When number of inactive bits is 0, result of the simulations for LNS and FP arithmetic are identical (Figure 4). Because of that next simulations uses only LNS arithmetic. Increasing of inactive bits (that
is decreasing of accuracy of LNS number) to 5 has no effect on the results of simulations, when LMS $\mu$ parameter is greater than 0.005 (Fig.4). It has effect, when $\mu = 0.001$, like in beginning of simulations and cause decreasing of disturbance attenuation (but only if tabulating of addiction and subtraction operations is active). It is follow from the above remarks, that interpolation algorithm is too simple and cause inaccuracy in arithmetical operations. Increasing of inactive bits at least to 15 , when LMS parameter $\mu = 0.01$ cause observable decreasing disturbance attenuation (Figure Fig.6). When number of inactive bits is 16 and more, ANC controller works improperly ($J < 0$).

5 Conclusions

In the paper logarithmic number system (LNS) in control oriented applications is presented. Both LNS and FP arithmetics were precisely presented, with a focus on representative of numbers, number formats, etc. Libraries for implement LNS in specific applications were developed. Both arithmetics were applied in a ANC application and results were compared. There were also behavior of ANC controller presented, when accuracy of used LNS arithmetic were decreased.

6 Acknowledgement

The financial support of this research (BW-404/Rau-1/2006) is gratefully acknowledged.

7 Reference

[3] Coleman, J; Chester, E.: A 32-bit Logarithmic Arithmetic Unit and Its Performance Compared To Floating Point
[8] Lee, B; Lever, K.: Logarithmic Number System and Floating-Point Implementations of a Well-conditioned RLS Estimation Algorithm on FPGA, University of Wales
Control algorithms for heating system in small building

Szymon Ogonowski
Silesian University of Technology, Institute of Automatic Control
Szymon.Ogonowski@polsl.pl

Abstract: The paper presents comparison of different methods of heating system control in small buildings. The comparison basis is constituted by the standard: boiler-, room- and weather-structured control. These simple algorithms cause temperature oscillations in the range that is tolerated by inhabitants. However, reductions of the oscillations amplitude could bring significant improvement if the fuel consumption is concerned. This became a motivation for searching of more sophisticated control algorithms. The paper presents cascade control structure with outer controller of PI and predictive type. The structure is also ameliorated by the static weather compensation. The algorithms are tuned to decrease temperature variation and to keep the mean value within the range provided by the standard controllers. Proper models were used for the algorithms synthesis. The models have been identified using four months observation of the real-world heating system.

1 Introduction

Technology of the heat sources especially used in a small building has been developing much faster than the corresponding control systems. Simple relay control dominates because dynamic of the plant (building) is significantly slower than dynamic of the heat source and imposed oscillations seems to be damped enough. The resulting temperature to be controlled oscillates with an amplitude acceptable by the user. However, from the fuel consumption viewpoint the problem is much different. Significant reduction of the fuel consumption can be achieved if the variation of the controlled temperature is attenuated [1]. To improve the heating control system performance advanced methods which orient the control system on the building dynamics have to be applied. To compare different types of control algorithms it is necessary to perform identical experiments. This is almost impossible because of the large number of factors influencing the plant behavior and other difficulties like necessity of keeping inhabitants’ thermal comfort or outer weather conditions strongly influence the plant. The experiment duration should be long enough to notice the plant reaction on different control algorithms being applied. On the other hand outer conditions can change in much shorter period. One winter period is usually too short to perform enough number of experiments. Conditions of winters changes from year to year what causes another limitations [8]. The modeling and simulation methods is then the only way to tackle the problem.

1.1 Data collection

The very first problem was the choice of the representative building for experiments. There are some examples of similar experiments in the literature [5, 6] but usually specially prepared enclosure and heat sources are used for data collection. In the case described in this paper the whole building is taken into account what enables struggling with real-environment problems. After careful consideration of all factors influencing thermal properties of the plant the typical building of 360 m² living space was chosen. The building is equipped with gas boiler with indoor control strategy. The controller and the building itself is described in [7]. To obtain proper model the identification data must be collected during normal use of the building. The system that collects data should not disturb inhabitants’ living conditions. In the same time it should register as many abnormal situations as possible. To perform all measurements special data collection system with wireless communication was created.
The system consists of two types of devices – Central Unit (CU) and Radio Module (RM/Server). Both equipped with: wireless communication, digital and analog I/O, 1-wire bus for digital thermo elements and RS-232. The general structure of the system is presented in Fig. 1. The system together with necessary data preprocessing software is detailed in [7]. During four months experiments the system measured following signals: 3 outdoor temperatures, 20 indoor temperatures (including boiler and hot water temperatures) and 7 system states (boiler and circulation pump gears and states and relays states).

![Fig. 1 The structure of data collection system](image)

### 1.2 Model identification

The plant dynamics is very complex because there are many heat sources in the building (boiler, fire place, oven etc.) and many ways of heat transfer (e.g. between boiler and radiators, radiators and rooms' interior, between neighbor rooms etc). The plant is strongly disturbed with outer conditions and the inhabitants behavior also causes temperature variations. The measurements were collected while the existing heating control system was on. This led to identification of the system with feedback (in our case with non-linear controller). Non-linear was also the plant itself and linear approximation was needed. Note that almost all the signals describing the plant state are continuous.

![Fig. 2 General structure of the heating system model](image)

The goal of the identification was to obtain model which can be useful for control algorithms synthesis. After assuming some simplifications the model structure was posed (Fig. 2). The general model consists of elementary models which explains the heat transfer in respective parts of the building. In the first approach all elementary models were identified with least square (LS) method as linear and stationary autoregressive exogenous models (ARX) [9]. Input signals of elementary models are continuous and not constant between sampling periods thus the identified discrete models were recalculated to continuous representation using Tustin method [10]. There were some problems with $K_h$ model identification because measured $T_h$ signal was disturbed. Another problem occurred while non-linear behavior of the plant was modeled. Solution of the problems and detail identification procedures are described in [13].
2 Simulations

To compare different control algorithms and structures 7-day period of collected data was chosen. During the period averaged outer temperature was varying from \(-16^\circ C\) to \(+7^\circ C\). Also other disturbances as kitchen or fire place temperature were changing as it is shown in Fig. 3. The simulated temperature set-point was different for the day and for the night similarly as in the real controller. Integrating the time when boiler is ON allows to calculate the fuel consumption. To validate model accuracy the calculated fuel consumption was compared with fuel-meter data. Error did not exceed 2% if the standard on-off controller was simulated (see Fig. 2).

2.1 Room control structure

After model validation three other controllers were simulated in room control structure. The first algorithm was the standard on-off with averaging feed-back [12]. The controller properties were similar to PD controller. The next two were PD controller and GPC controller with two state output. The regulators’ parameters were found using Gauss-Seidel optimization. Simulation results are presented bellow:
It can be seen the temperature variation is reduced. It is possible then to level down temperature day set-point from 22.4°C to 22.2°C. For inhabitants this change is not noticeable (temperature still will be varying between 22.2°C and 22.7°C) but the fuel consumption can be reduced what is the most important result.

Table 1  Gas consumption for room control structure

<table>
<thead>
<tr>
<th>Set-point [°C]</th>
<th>ON/OFF [m³]</th>
<th>ON/OFF with aver. f-b [m³]</th>
<th>PD [m³]</th>
<th>GPC [m³]</th>
</tr>
</thead>
<tbody>
<tr>
<td>22.4</td>
<td>176.7</td>
<td>173.6</td>
<td>176.5</td>
<td>174.8</td>
</tr>
<tr>
<td>22.2</td>
<td>141.9</td>
<td>139.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

One can suspect, that limitation of the relay hysteresis can give similar results as presented above. Such tests were also performed and the results were similar only if no constraints on switching frequency were imposed. Very small hysteresis caused fast relay switching – ON state duration was often shorten then a few seconds. The boiler needs about 5 sec to light up. Very fast boiler switching is then not economic from the fuel consumption viewpoint. Relay with small hysteresis leaves no room for precise switching time control.

2.2 Cascade structure

It is possible to extract two parts of the plant: boiler with hot water circulation system and the building with radiators (heat consumption). The cascade structure can then be applied with inner feed-back from the hot water temperature (boiler output) and outer feed-back from the hall temperature. The structure is presented in Fig. 5a. In our case the boiler has two-state input, so simple two-state relay is proposed as the inner controller. Good tracking properties are required for the outer controller so PI and GPC controllers were tested. The simulation results for both controllers are similar – the fuel consumption and temperature oscillations were reduced while comparing with the room control structure.

Table 2  Gas consumption comparison

<table>
<thead>
<tr>
<th></th>
<th>without cascade [m³]</th>
<th>with cascade [m³]</th>
</tr>
</thead>
<tbody>
<tr>
<td>PI</td>
<td>141.9</td>
<td>134.3</td>
</tr>
<tr>
<td>GPC</td>
<td>139.5</td>
<td>134.0</td>
</tr>
</tbody>
</table>
2.3 Weather compensation

The most complex of commercial heating system controllers is weather compensation approach (see Fig. 7). A user has to choose so called heating curve [2] which is nothing else but the compensator gain choice. The most important disadvantage is neglecting of the building dynamical properties – only a steady state is taken into account. The consequence is the fuel over-consumption together with under and overheating [3]. These effects were observed during simulations if standard commercial heating curves were used. It was a motivation of identification of model-dedicated heating curves. Data was obtained by simulation of the plant controlled by PI controller as shown in Fig. 6a. During one simulation $T_o$ and $T_{set}$ signals were constant and $T_h$ in steady-state was observed. This allowed to design heating surface for the model (Fig 6b). One can obtain standard heating curve from heating surface by choosing desirable value of $T_{set}$. Two controllers were tested in this structure: standard ON/OFF and PI. Results obtained during simulation with dedicated heating curves are presented in Table 3.

![Fig. 6a Structure for heating curves identification (a) and heating surface (b)](image)

**Table 3  Gas consumption comparison**

<table>
<thead>
<tr>
<th>Set-point</th>
<th>ON/OFF</th>
<th>PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>22.4 °C</td>
<td>204 [m³]</td>
<td>191.7 [m³]</td>
</tr>
<tr>
<td>22.2 °C</td>
<td>162.1 [m³]</td>
<td>161.0 [m³]</td>
</tr>
</tbody>
</table>

Note that temperature set-point for this structure was constant for day and night. This follows from the fact, that there is no evident feed-back from $T_h$ signal. Such simulations enables comparison with real-environment experiments’ results.

![Fig. 7 Weather compensation structure](image)

**2.4 Cascade structure with static compensation**

After analysis of presented results more complex control structure is proposed namely the cascade structure with PI as the outer controller. The structure is ameliorated by adding static compensation of outer temperature (heating curves). For comparison with standard weather compensation approach day and night temperature set-points were equal. Hot water temperature (determined by the outer controller) is corrected according to the outer temperature (see Fig. 8a). The simulation results are presented in Fig 8b. The fuel consumption for this structure was reduced to 153.2m³ if the 22.2°C set-point was simulated.
3 Conclusions

The paper presents first approach for modeling and simulations method of control algorithms synthesis. The simulations results proved accuracy of the obtained model and showed possibilities of fuel consumption reduction if advanced control algorithms are used. Presented cascade structure with static weather compensation will be the basis for further researches of multi-level control strategies for heating systems.

4 References

Direct Torque Control with Reduced Torque Ripple for Permanent Magnet Synchronous Motor

Sanda Victorinne Paturca, Paula Deaconescu
University Politehnica of Bucharest
sanda@amotion.pub.ro, paula_d@amotion.pub.ro

Abstract: This paper proposes a method of reducing the torque ripple for Direct Torque Control (DTC) of a permanent magnet synchronous motor (PMSM) drive. In conventional Direct Torque Control, a single stator voltage vector of the inverter standard topology is selected every control sampling period, and it is maintained for the whole period. By this switching technique, based on hysteresis, large and small torque error are not differentiated, which cause an extra torque ripple in motor steady state operation. To overcome this problem, this paper proposes a solution, which consists in the modulation of the nonzero voltage vector duration over a sampling period, according to the instant value of the torque error. For this purpose it was introduced a duty ratio proportional to the torque error. The presented results show the torque ripple reduction obtained by using the proposed method. Its main advantage is that it requires an insignificant additional computation, preserving the simplicity of the conventional DTC.

1 Introduction

In 1996, DTC was introduced on the market by ABB [2], which consider it a viable alternative to Vector Flux Oriented Control (FOC). The main advantages of DTC are the simple control scheme, a very good torque dynamic response, as well as the fact that it does not need the rotor speed or position to realize the torque and flux control (for this reason DTC is considered a "sensorless" control strategy). These advantages can be fully exploited in those electric drives where not the speed, but only the torque is to be controlled. For this kind of applications, DTC can be a very attractive option, because it is able to provide high dynamic performance at convenient costs.

However, the classical DTC has some drawbacks, and one of these is the significant torque and current ripple generated in steady state operation. Taking into account the large slopes of the resulted torque and the fact that only a single voltage vector is applied to the inverter in a control sampling period, the classical DTC needs high sampling frequencies (above 40 kHz [2]) to obtain a good steady state behavior. This requires high performance controllers, like the DSP, which rises the overall cost of the drive.

This paper proposes a simple solution for reducing the torque ripple in classical Direct Torque Control, while preserving the good dynamic and structural simplicity of this scheme. The proposed method consists in the modulation of the nonzero voltage vector duration over a sampling period, according to the torque and stator flux errors. The algorithm for duty ratio calculation requires only a few arithmetic operations and is very easy to implement on a fixed point digital processor. There results are presented for several motor operating points, to show the improved steady state operation as compared to conventional DTC.

2 Direct Torque Control

The basic model of the classical DTC PMSM motor scheme is shown in Fig. 1. It consists of torque and stator flux estimators, torque and flux hysteresis comparators, a switching table and a voltage source inverter (VSI). The configuration is much simpler than that of the FOC system where frame transformation, rotor position or speed sensors are required. The basic
The idea of DTC is to choose the best voltage vector in order to control both stator flux and electromagnetic torque of machine simultaneously [1]. At each sample time, the two stator currents $i_{SA}$ and $i_{SB}$ and DC-bus voltage $U_{DC}$ are sampled. The components of the stator voltage space vector in the stationary reference frame are calculated as shown in (1) and (2) [3].

$$u_{\alpha} = \frac{2}{3} U_{DC} \left( S_A - \frac{S_B - S_C}{2} \right)$$

$$u_{\beta} = \frac{2}{3} U_{DC} \frac{S_B - S_C}{\sqrt{3}}$$

where: $S_A$, $S_B$, $S_C$ denote the inverter switching states, in which $S_i = 1$ ($i = A, B, C$), if the upper leg switch is on and $S_i = 0$, if the upper leg switch is off;

The α–β components of the stator current space vector are calculated using equations (3) and (4), supposing the motor has the star connection.

$$i_{\alpha} = i_{A}$$

$$i_{\beta} = \frac{i_{A} + 2i_{B}}{\sqrt{3}}$$

Using the equations (1)–(4) and the stator resistance, the α–β components of the stator flux are calculated in (5) and (6):

$$\psi_{\alpha} = \int (u_{\alpha} - R_s i_{\alpha}) dt$$

$$\psi_{\beta} = \int (u_{\beta} - R_s i_{\beta}) dt$$

The circular trajectory of stator flux is divided into six symmetrical sections ($S_1$ – $S_6$) referred to inverter voltage vectors, as shown in Figure 2. The α–β components of the stator flux are used to determine the sector in which the flux vector are located.

Then using equations (3) – (6), the magnitude of the stator flux and electromagnetic torque are calculated in (7) and (8).

$$|\psi| = \sqrt{\psi_{\alpha}^2 + \psi_{\beta}^2}$$

$$T_e = \frac{3}{2} P (\psi_{\alpha} i_{\beta} - \psi_{\beta} i_{\alpha})$$

where: $P$ is the number of pole pairs

$R_s$ is the Stator resistance

The calculated magnitude of stator flux and electric torque are compared with their reference values in their corresponding hysteresis comparators as are shown in Fig. 1. Finally, the outputs of the comparators and the number of sector at which the stator flux space vector is
located are fed to a switching table to select an appropriate inverter voltage vector [3]. As shown in Fig. 2, eight switching combinations can be selected in a two-level voltage source inverter, two of which determine zero voltage vectors and the others generate six equally spaced voltage vectors having the same amplitude.

![Inverter voltage vectors and stator flux sectors](image)

The selected voltage vector will be applied to the PMSM motor at the end of the sample time.

### 2.1 The proposed method for torque ripple reduction

In conventional DTC the voltage vector selection is based on the torque and flux errors, but small and large errors are not distinguished by the hysteresis controllers. The voltage vectors are applied for the entire sample period, even for small errors, resulting in large torque overshoots in steady-state regime.

The method for torque ripple reduction, presented in this paper, consists in the modulation of the nonzero voltage vector duration over a sampling period, according to the torque error, using symmetric PWM. The nonzero voltage vectors are selected using a switching table, like in classical DTC. The null vectors are automatically inserted by using PWM, so they are no more needed in the switching table. Consequently, in the proposed control scheme, shown in Fig. 3, simple comparators, with no hysteresis, are used. A duty ratio calculator and a PWM block have been added to the classical scheme.

![Block diagram of the proposed control](image)

The calculation procedure of the duty ratio is presented below.

In order to preserve the very good dynamic response of the classical DTC, the proposed method for reducing the torque ripple is applied only when the actual torque value is located in a "proximity" zone, around the reference value. This zone, whose width is denoted by $Z_T$,
is similar to the torque hysteresis band of the classical DTC torque comparator. In other words, only for torque errors between $-Z_T$ and $Z_T$ the voltage duty ratio is modified, otherwise it is kept at 100%. The duty ratio was taken equal to the normalized torque error – which means the ratio of the torque error, $e_T$, to the zone width, $Z_T$ – as shown in (9).

$$\delta_T = \left| \frac{e_T}{Z_T} \right| \leq 1$$

where: $\frac{e_T}{Z_T}$ is the normalized value of the torque error

$\delta_T$ is the duty ratio corresponding to the torque error.

In Fig. 4 is represented the duty ratio as function of normalized torque error $\left| \frac{e_T}{Z_T} \right|$.

The duty ratio $\delta$ is limited to a minimum value, $\delta_{\text{min}}$, in order to consider the maximum inverter switching frequency.

It should be noted that no proportional coefficients were introduced, since the slope of the duty ratio characteristic can be adjusted by simply resizing $Z_T$.

![Fig. 4 The characteristics of duty ratio (\(\delta\))]{}

Every sample time, one of the six nonzero voltage vectors are selected from the switching table, and it is applied to the inverter using symmetric PWM switching strategy. Considering the calculated duty ratio, the selected voltage vector, given by $(S_A, S_B, S_C)$ is applied to the inverter in the following manner: if $S_i = 0$, the pulse duty cycle of the corresponding inverter leg is zero, otherwise it is $\delta \cdot \text{PWM period}$.

The line voltage at the inverter output is given by (10).

$$U_I = \delta \cdot U_{DC}$$

where: $U_I$ is the line voltage at the inverter output

$U_{DC}$ is DC link voltage.

At current sample time, the $\alpha - \beta$ components of the stator voltage are calculated using (11) and (12), based on (1) and (2) respectively.

$$u_{sa} = \frac{2}{3} \delta U_{DC} \left( S_A - \frac{S_B - S_C}{2} \right)$$

$$u_{sb} = \frac{2}{3} \delta U_{DC} \frac{S_B - S_C}{\sqrt{3}}$$

where $\delta$ is the duty ratio calculated at previous sample time.

### 2.2 Results

The simulations were made in Matlab-Simulink environment, for a PMSM with nonsalient poles. Graphical results are presented for few operating points, showing the motor no load operation, followed by the operation with rated torque, which is applied as step at $t=0.2s$. The
control sampling period was set to 100 µs. The PWM period is considered equal to the control sample period. The results obtained by conventional DTC are represented in Fig. 5 and 6, for low and respectively high frequency.

In Figure 7 and 8 are shown the results obtained by DTC using the proposed solution for torque ripple reduction, for low and respectively high frequency. The voltage duty ratio is related to the torque error, according to equation (9). There can be seen the diminished torque ripple as compared to conventional DTC.

2.3 Conclusions
This paper presents a cost-effective solution for reducing the torque ripple in conventional Direct Torque Control, while preserving the dynamic response and structural simplicity of this control scheme.
The solution proposed uses PWM control of the voltage vector with duty ratio determined using a simple equation based on the torque error. The results show that the torque ripple was reduced. The proposed method does not require additional motor parameters or knowledge of the rotor speed.

Taking into account that today many microcontrollers and DSP offer PWM interfaces, the proposed method can be very easily implemented on them, adding very little computation effort to the classical DTC control algorithm.

3 References


Abstract: This presentation introduces a portable system created for measurements of vectorial effects in sound intensity field. The presented system enables the realization of the serial sound intensity measurements and consequently, also the determination of the spatial vector distribution of sound intensity. To estimate sound intensity vector, its two components - sound pressure and sound particle velocity must be defined.

For the measurement of the sound pressure, the condenser microphone is used. For the measurement of the sound particle velocity, three Microflown sensors are used. Note that sound pressure is a scalar and particle velocity is a vector quantity. The result of multiplying sound pressure and each of the acoustic particle velocity is the sound intensity vector.

1 Sound intensity measurement

Sound intensity \( I \) is the average rate at which sound energy is transmitted through a unit area perpendicular to the specified direction at the point considered. It is a vector quantity defined as the time averaged product of the sound pressure \( p \) (scalar) and the corresponding particle velocity \( u \) (vector) at the same position.

\[
I \left[ \frac{W}{m^2} \right] = \lim_{T \to \infty} \frac{1}{T} \int_{0}^{T} \hat{u}(t) \cdot p(t) dt
\]  

(1)

The frequency distribution of the mean intensity is

\[
I(\omega) = Re\{G_{pu}(\omega)\}
\]  

(2)

where \( G_{pu} \) is cross-spectrum of the sound pressure and the particle velocity signals [1].

For the measurement of the sound intensity, the USP Microflown probe is used. USP - the Ultimate Sound Probe, it is an ultra miniature three dimensional sound probe that is capable of measuring broad banded sound pressure and particle velocity in three orthogonal directions [4].

The USP is a compact and fully integrated sound probe that combines three orthogonally positioned particle velocity sensors and a miniature pressure microphone. The sensor configuration without its cap is less than 5 mm × 5 mm × 5 mm.
To take a full advantage of the Microflown probe it’s necessary to use 4-channels analyzer. Applied instrument should enable calculation of auto-spectrum and cross-channel functions like transfer function or cross-spectrum. Additional requirement is possibility of control of the analyzer from external application so that enables the realization of the serial sound intensity measurements and consequently, also the determination of the spatial vector distribution of sound intensity.

All requirements are performed by the Siglab type 20-42 [3]. It’s 4-channels analyzer and generator. Software dedicated to control of Siglab is divide in a few programs. Each one of them makes possible executing of specific operation, for example:

- The Hardware Performance Verification (HWVERIFY) – calibration of Siglab’s input and output channels
- Dynamic Signal Analyzer (VNA) – analyse of signals
- Virtual Third Octave (VTO) – full-octave and fractional-octave measurements
- Function Generator (VFG) – generation of signals

Described above construction of software make difficult to connect of they capabilities, especially when control of the system from external application is required. The most important parts of Siglab’s software in sound intensity measurement is Dynamic Signal Analyzer (VNA) and Virtual Third Octave (VTO). First of them is use to signals record and cross-spectrum calculation. Second of the is use to estimation of sound intensity value in fractional-octave bands. Direction for use of VNA and VTO in the measurement system is describe on fig 3. In brackets defined which part of the system execute specified task. By “APP” defined program writed by author.

Scheme on fig 3 demonstrate the most important steps in the measure and calculations of the sound intensity.

![Fig. 1 The Microflown USP probe](image1.png)

![Fig. 2 The sound intensity measurement circuit](image2.png)
Presented system makes possible three types of measurement:

- Three-dimensional sound intensity measurement in 1/12 octave bands
- Three-dimensional sound intensity measurement in 1/3 octave bands
- One-dimensional sound particle velocity measurement in 1/12 octave bands.
Important parts of computations of the sound intensity

2.1 Narrowband sound intensity measurement

Narrowband sound intensity measurement is the first step in estimation of the sound intensity value in fractional-octave bands. Measurements are made using a Dynamic Signal Analyzer (VNA). A scheme in figure 5 illustrates this part of the algorithm. Cross-spectrum estimated in this step is the complex value. For the next operations only the real part of it is important.

Fig. 5 Scheme of narrowband sound intensity measurement

2.2 Filtering of the sound intensity with Virtual Third Octave (VTO)

Because VTO and VNA do not use the same type of data file, information about signals must be transferred from VNA’s data file to VTO’s data file. When the file with replaced data is opened in VTO, fractional-octave bands values of signals are calculated automatically. Note that the used file must be prepared earlier. Disadvantage of VTO in this application is that VTO calculates only RMS (unsigned) value of signals in bands, therefore the sign of the vector must be estimated in the next step of the algorithm. This is executed by integrating the sound intensity in bands.

Fig. 6 Scheme of fractional-octave bands filtering
3 Testing of system

Created system was tested and calibrated by compare results with results acquired on system constructed with use of Norsonic RTA 840 [2]. RTA 840 it’s two channel analyzer adapted for one-dimensional sound intensity measurement. Differences between them are smaller then 0.3dB. Comparison of results is presented on fig 7.

![Fig. 7 Results of the sound intensity field measurements on Siglab and RTA 840](image)

4 Summary

Presented system make the three-dimensional sound intensity measurement three time faster than measurement realized on system based on RTA 840. This is very important in measurements of vectorial effects in sound intensity field because to estimation and analysis of vectorial effects in sound intensity field thousands of sound intensity measurements must be done.

![Fig. 8 Flux of the sound intensity field measured on open end of the standing wave tube](image)

Fig 8 illustrate flux of the sound intensity field measured on open end of the standing wave tube. It’s very simple example of the sound intensity field measurement made only for testing system. Examples of using the sound intensity measurement technique and visualisation of the vectorial effects in the sound intensity field may be found in many publication wrote by dr hab. inż head of Applied Vibroacoustic Dep in Maritime Faculty of Technology in Szczecin University of Technology.
5 References


Identifying an electric oven with two heating zones

Elena Ioana Vățăjelu
Technical University of Cluj Napoca, România
Ioana.Vatăjelu@aut.utcluj.ro

Abstract: Obtaining a constant temperature or a certain temperature gradient on a surface or in a volume is a very important and difficult control task. Our research is focused in this direction and the first step is to develop a model of the heating process. The paper describes a method to determine a mathematical model for an electric oven with silicon carbide heaters, considering it has two heating zones. The model will be obtained through experimental methods using the heating curves. Once the transfer function for each zone and the interinfluence transfer functions are determined, the transfer matrix can be written. Based on this model, the further work will consist in applying some control strategies on this process in order to achieve a certain temperature gradient on a surface and then choosing the most suited one and trying to improve it.

1 Introduction

Ovens are used in many fields, such as pottery, metallurgy, electronics, medicine, chemistry, where thermal processing and treatment is needed. In order to obtain the temperature variation required for a specific application, the oven must be controlled; therefore its dynamic behavior must be known.

Usually, static and dynamical characteristics are experimentally determined on the real oven, by applying a step or sinusoidal input signal. This oven was previously identified considering it has only one heating zone [1]. The paper describes the test bed setup, the main characteristics of its components and a method to identify the system taking in account it has two heating zones. Some assumptions will be made: that there are no leaks and the heat is transferred to the exterior only through the walls by conductivity. Another assumption was that the power voltage is stabile and that each heater is in the center of the corresponding heating zone.

2 Test Bed Setup

The experiments were performed on an electric oven with silicon carbide heaters. The heat is produced by four SiC heating elements (two by two in parallel), supplied via the power drive. The temperature sensors are two identical type K thermocouples, connected to a network data acquisition and control module. The data is transmitted to the computer through the serial port and memorized (fig. 1).

Fig. 1 The block diagram
2.1 The oven
The oven has the capacity of 0.02 m$^3$ and the nominal power is 6kW. The maximum temperature that can be achieved in this oven is 1300°C. The dynamic behavior of the oven is mainly influenced by the capacity of the working space, the power of the heater, the nature of the firebrick and the type and placement of the temperature sensors.

The walls of the oven are made of refractory brick, which usually has the highest thermal conductivity (for chamotte - K=0.84W/mK [2]). The heating element converts electricity into heat through the process of Joule heating. Silicon Carbide heating element is a non-metal electric heater. When it is heated in air by electricity until surface temperature reaches $1050\pm50^\circ\text{C}$, its resistivity will be in the range of 600-1400 $\Omega\text{mm}^2/\text{m}$ [4] (in this particular application the resistance of the heater is $9\,\Omega$ and its resistivity is 706.85 $\Omega\text{mm}^2/\text{m}$ at room temperature). The resistance of SiC heater changes with increasing temperature. From room temperature to 800, the curve of resistance-temperature is negative and turns positive when the temperature is above 800°C [7].

The temperature is measured with two identical type K thermocouples (Chromel (+), Alumel (-)). They offer excellent accuracy over a wide temperature range (from 95 to 1250 °C). Thermocouples are based on the principle that when two dissimilar metals are joined a predictable voltage will be generated that relates to the difference in temperature between the measuring junction and the reference junction. The thermocouples used in this application have a narrower operating range; they can measure temperatures up to 550°C.

2.2 The Power Drive
The heat is produced by four SiC heating elements (two by two in parallel), each group supplied individually via the power drive. The power drive is based on two WG 480 D10 Z solid-state relays. Using solid-state devices, there are no moving parts to wear out, and they are able to switch on and off much faster than any mechanical relay armature can move. There is no sparking between contacts, and no problems with contact corrosion.

The WG 480 D10 Z’s features are: switching – zero cross, output – back-to-back SCR with internal snubber, input – DC with constant current control. One significant advantage of a solid-state SCR relay over an electromechanical device is its natural tendency to open the AC circuit only at a point of zero load current [4]. The circuit will never be interrupted in the middle of a sine wave peak. Such untimely interruptions in a circuit containing substantial inductance would normally produce large voltage spikes due to the sudden magnetic field collapse around the inductance. This will not happen in a circuit broken by an SCR. This feature is called zero-crossover switching. The output signal of the solid-state relay is applied on each heating group through a 7000 DIO ICP-CON module that outputs a PWM voltage with a variable duty cycle. (The DIO modules support TTL signal, relay contact outputs, solid-state relay output, and open-collector output [3]).

2.3 Data Acquisition
The oven temperature is actually the air temperature in the working space and it is measured with two thermocouples. The sensors are directly connected to the network data acquisition and control module (ICP CON I – 7019). The I – 7019 is an 8 – channel voltage, current, and thermocouple input module, with the ability to connect various types of inputs to a single module [3]. This module interfaces with the computer by serial port through the I-7024 module of the same family (ICP-CON) and the data from the sensor is saved, processed and plotted using Matlab environment.
The block diagram of the process is represented in figure 2, where the values of the step signals are given by the duty cycle of the PWM (for each group individually). The temperature is memorized and plotted on the scope.

Fig. 2 The block diagram in Simulink

3 Results

In order to work with an oven one must know its dynamic behavior, which depends on the capacity of the working space, the power of the heating elements and the nature of the firebrick. Considering that the oven is small and has a single layer of firebrick it can even be approximated as a first order system (this is the case for this particular oven [1]). Because the main objective is to obtain a constant temperature or a certain temperature gradient on a surface (or in a volume) using this oven, the working space was divided in two zones. To write the differential equation that will describe the thermal behavior for each zone is some how difficult and inexact, so an experimental identification is more appropriate.

The experiments presented below took place on a laboratory oven (the one described above), the duty cycle was set at 30% (which will be considered as the unit step response). The step was applied successively on the heating groups and the temperatures for each thermocouple were memorized. So when the step was applied on the first heating group the curves from figure 3a were obtained and when the step input was applied on the second heating group the curves from figure 3b were obtained.

Fig. 3 – a – temperature versus time when the first heating group is on 
– b – temperature versus time when the second heating group is on

From these two diagrams it is easily observed that the two zones do not have the same dynamical behavior. Some of the reasons may be that the heaters are worn-out differently and the temperature loss through the walls is not uniform. Considering this, an experimentally determined model will be more precise than an analytical one.

It is obvious, from the experimental data that this oven behaves like a first order system. A direct search method of optimization that works well for stochastic problems was used.
(Nelder – Mead method). This method is based on evaluating a function at the vertices of a simplex, then iteratively shrinking the simplex as better points are found until some desired bound is obtained. For two variables, a simplex is a triangle, and the method is a pattern search that compares function values at the three vertices of a triangle. The algorithm is stated using the term simplex (a generalized triangle in n dimensions) and will find the minimum of a function of n variables. The method is effective and computationally compact [9].

So, starting with the time variant equation [8]:

\[ y(t) = 1 - ke^{pt} \quad (1) \]

that defines the first order system response to a step input, with the Nelder-Mead method, the parameters k and p of the equation that better matches the experimental data, will be found (the experimental data must first be normalized). The figure 4 represents the simulation results.

For the first thermocouple k=1 and p= -5.7778 \times 10^{-5}, so (1) becomes:

\[ y(t) = 1 - e^{-5.7778 \times 10^{-5} t} \quad (2) \]

and for the second thermocouple k=1 and p= -5.0125 \times 10^{-5}, so (1) becomes:

\[ y(t) = 1 - e^{-5.0125 \times 10^{-5} t} \quad (3) \]

The transfer function of the system that has the step response defined by (2) is:

\[ H_{11}(s) = \frac{1}{1.731 \times 10^4 s + 1} \quad (4) \]

and the transfer function of the system that has the step response defined by (3) is:

\[ H_{12}(s) = \frac{1}{1.995 \times 10^4 s + 1} \quad (5) \]

The same procedure is applied when the second heating group is on. The figure 5 represents the simulation results.
For the first thermocouple \(k=1\) and \(p = -5.1647 \times 10^{-5}\), so (1) becomes:
\[
y(t) = 1 - e^{-5.1647 \times 10^{-5} t}
\]
and for the second thermocouple \(k=1\) and \(p = -6.637 \times 10^{-5}\), so (1) becomes:
\[
y(t) = 1 - e^{-6.637 \times 10^{-5} t}
\]
The transfer function of the system that has the step response defined by (6) is:
\[
H_{21}(s) = \frac{1}{1.936 \times 10^4 s + 1}
\]
and the transfer function of the system that has the step response defined by (7) is:
\[
H_{22}(s) = \frac{1}{1.507 \times 10^4 s + 1}
\]
Defining the step signal applied to the 1st heating group as \(u_1\), the step signal applied to the 2nd heating group as \(u_2\), the temperature obtained with the 1st thermocouple as \(y_1\) and the temperature obtained with the 2nd thermocouple as \(y_2\), the system can be represented like in figure 6. \(H_{11}(s)\) and \(H_{22}(s)\) are the direct transfer functions and \(H_{12}(s)\) and \(H_{21}(s)\) are the interinfluence transfer functions.

\[
\begin{pmatrix}
y_1(s) \\
y_2(s)
\end{pmatrix} = G(s) \begin{pmatrix}
u_1(s) \\
u_2(s)
\end{pmatrix}
\]

The transfer matrix for the MIMO system represented above is:
\[
G(s) = \begin{bmatrix}
H_{11}(s) & H_{21}(s) \\
H_{12}(s) & H_{22}(s)
\end{bmatrix}
\]

\[
\begin{pmatrix}
y_1(s) \\
y_2(s)
\end{pmatrix} = \begin{pmatrix}
1 & 1 \\
1.731 \times 10^4 s + 1 & 1.936 \times 10^4 s + 1 \\
1.995 \times 10^4 s + 1 & 1.507 \times 10^4 s + 1
\end{pmatrix} \begin{pmatrix}
u_1(s) \\
u_2(s)
\end{pmatrix}
\]
4 Conclusions

The paper describes the test bed setup, the main characteristics of its components and a method to identify the system taking in account it has two heating zones. The heating process was considered a linear one, even though it isn’t (the oven’s dynamics is piecewise linear). The simplest model was obtained (a first order one) with a good match between experimental data and the simulated step response.

The most important step in controlling a system is knowing its dynamic and static behavior. The current results will be further used for designing a control strategy for this system. Heating in this oven an irregular aluminum object we will try to obtain a constant temperature on its surface using different control strategies and than comparing them.

5 References

Section 2

Electronics
Fault diagnosis of analog electronic circuits

Łukasz Chruszczyk (lchruszczyk@polsl.pl)
Damian Grzechca (dgrzechca@polsl.pl)
Jerzy Rutkowski (jrutkowski@polsl.pl)
Silesian University of Technology
Gliwice, Poland

Abstract: this paper presents combined approach to testing of analog electronic circuits using evolutionary algorithms, information theory and wavelet analysis. Testing procedure is based on specialised aperiodic excitation that maximises probability of fault detection and location. Results are compared with testing using unit step excitation. The testing method belongs to SBT (Simulation Before Test) class of fault testing procedures and focuses on most difficult case where very few (usually only input and output) nodes of integrated circuit under test (CUT) are available.

1 Introduction

Fault analysis of analog electronic circuits is a difficult task, much more complex than testing of digital circuits. Spread of values of components (caused by design) is the greatest problem. Fault diagnosis of analog electronic circuits can be performed by means of three types of signals.

DC testing is the simplest method. Testing procedure is based on measurement of voltages at selected nodes and currents in selected branches. Simplicity of measuring instruments and fast testing time are advantages. The method has some limitations too. Except parametric and catastrophic testing of resistive elements, it can only detect some catastrophic faults of energy storage elements (short capacitors and disconnected inductors).

AC testing is second method. Single- or multi-frequency signal is applied to circuit, then the output is analysed (magnitude and spectrum of appropriate frequency components). The method can be applied to circuits containing both resistive and energy storage components. Disadvantages are more complex measurement instruments and longer testing time as compared to DC testing.

Third method is testing by means of specialised aperiodic excitations. This method is the most complex, but in many cases performs much better than DC and AC testing. Response of circuit for aperiodic excitation may deliver more information about CUT, which is not always achievable with the AC excitation (especially when only a few frequencies are used). Testing using aperiodic excitation usually performs better than DC and AC testing when access to circuit internal nodes is limited. This is often case of integrated circuits (IC), where generally, only input, output and supply current nodes (from power supply and to ground) are accessible. The DC and AC testing may be helpless in such situation. Longer testing time and complexity of measurement instruments are main disadvantages.

Complex aperiodic excitation may improve testability. The signal is usually applied to CUT input terminal. It is also possible to excite CUT in other ways e.g. modify supply voltage. Afterwards, CUT response (as well as supply current variations) are analysed in time or frequency domain. Testing by means of specialised signals is still the least popular comparing to DC and AC testing methods. However, it is very promising. A question arises: what should be a shape of the excitation signal to achieve the greatest possible fault detection and location?
2 Description

This paper presents a method that finds optimal or suboptimal shape of aperiodic excitation. Such signal maximises probability of fault location and detection. The excitation is encoded as vector of digital samples. Easy signal creation by means of D/A converter is the advantage. Selection of excitation sampling period is very important. Too long time causes that CUT treats input signal as sequence of step functions – not as approximation of analog signal. Too short sampling time (over-sampling) is also undesired. The excitation signal in its digitised form contains unnecessarily many samples. A rule of thumb is to set sampling time of the excitation not greater than 10% of CUT shortest time constant. Chosen value is close to the shortest time constant of the CUT, which is sufficient compromise between amount of data to process and approximation quality of analog excitation. The procedure uses:

a) genetic algorithm (GA) as “search engine”
b) information theory as descriptor of fault location and detection efficiency
c) wavelet transform as “feature extractor”
d) fault dictionary used to “manage” defined states (signatures) of the CUT.

2.1 Step 1 – initialisation of the GA

The GA is used as “search engine” to find values of samples of the excitation signal that maximises probability of fault detection and location. There is a population of individuals, each contains excitation signal encoded as vector of digital samples. Values of the samples are encoded as floating point numbers. Initial shape of signal held by the population is the unit step.

2.2 Step 2 – calculation of the signatures

There are calculated CUT responses (signatures) for all defined faults and healthy circuit. The signatures are used to build fault dictionary which is used in process of classification of CUT in unknown state [4].

2.3 Step 3 – CUT testing and classification

There is number of circuits generated in order to evaluate found solution. There are two parametric faults defined for each element: above and below its nominal value. Each parametric fault \( E_i \) is modelled as a change of nominal value by twice of its tolerance \( t_E \):

\[
E_i = E_{\text{nom}} \cdot (1 \pm 2 \cdot t_E) \quad (1)
\]

Values of non-faulty elements are chosen randomly (uniformly) within their tolerance range. Classification of CUT in unknown state is based on distance of its response from stored signatures in fault dictionary. The smallest distance criterion is used in all cases. The distance is calculated as Euclidean distance between response of CUT \( s_i \) and the stored signature \( s_r \):

\[
d = \sqrt{\sum (s_i - s_r)^2} \quad (2)
\]

Additionally wavelet transform is used to enhance differences between responses of healthy and faulty CUT. The difference between wavelet coefficients of analysed responses is computed using Euclidean distance as well. Utilisation of spectral as well as time domain information (impossible for Fourier transform) is great advantage of wavelet transform [1]. The transform is able to bring out features of analysed signal that can be difficult to detect by means of separate time or frequency analysis.

The GA may also analyse a frequency spectrum of found excitation. This promotes solutions with narrower bandwidth (less high frequency components). The spectrum is divided into two
equal bands: \( \text{low} = [f_0; f_{\text{int}}) \) and \( \text{high} = (f_{\text{int}}; f_{\text{max}}) \). Afterwards, fitness value \( \text{fit} \) of the given solution is modified, according to power spectral density in appropriate band \((P_{\text{low}}, P_{\text{high}})\):

\[
\text{fit} = \begin{cases} 
\text{fit} \cdot 2 & \text{if } P_{\text{low}} > P_{\text{high}} \\
\text{fit} \cdot 0.5 & \text{if } P_{\text{low}} \leq P_{\text{high}}
\end{cases}
\]  

(3)

2.4 Step 4 – evaluation of found solution

Information theory is utilised in order to evaluate efficiency of circuit states separation. Model of testing process is based on two-symbol information source and lossy information channel (fig.3), where \( p_{ij} \) is probability of symbol \( j \) reception, if symbol \( i \) has been transmitted [3].

![Fig.1 Binary information channel](image1)

![Fig.2 Schematic of the low pass filter](image2)

The probabilities \( p_{ij} \) are calculated as ratio of number of circuits classified to state \( D_j \) to number of circuits in state \( F_i \):

\[
p_{ij} = \frac{D_j}{F_i}
\]  

(4)

The probabilities \( p_{ij} \) can be written in matrix of fault location probabilities:

\[
\begin{array}{cccc}
D_0 & D_1 & D_2 & \ldots & D_N \\
F_0 & P_{00} & P_{01} & P_{02} & \ldots & P_{0N} \\
F_1 & P_{10} & P_{11} & P_{12} & \ldots & P_{1N} \\
F_2 & P_{20} & P_{21} & P_{22} & \ldots & P_{2N} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
F_N & P_{N0} & P_{N1} & P_{N2} & \ldots & P_{NN}
\end{array}
\]

where:

- \( F_0 \) – healthy circuit
- \( F_i \) – faulty circuit in \( i^{\text{th}} \) state
- \( D_i \) – decision: circuit in \( i^{\text{th}} \) state
- \( P_{ij} \) – probability that circuit in \( i^{\text{th}} \) state is classified to \( j^{\text{th}} \) state

Main diagonal of the matrix contains probabilities of correct classifications. Used optimisation criterion is to obtain values at main diagonal as close 1 as possible and values outside main diagonal as close 0 as possible.

In case of fault detection (test GO/NO GO), probabilities of fault detection can be written as follows:

\[
\begin{array}{ccc}
D_0 & D_f \\
H & P_{hh} & P_{hf} \\
F & P_{fh} & P_{ff}
\end{array}
\]
where:
- \( H \) – healthy circuit
- \( F \) - faulty circuit
- \( D_h \) – decision: circuit healthy
- \( D_f \) – decision: circuit faulty
- \( P_{hh} \) – probability of correct classification of healthy circuit
- \( P_{ff} \) – probability of correct classification of faulty circuit
- \( P_{fh} \) – probability of classification of faulty circuit as healthy one
- \( P_{hf} \) – probability of classification of healthy circuit as faulty one

The probabilities are computed from matrix of fault location probabilities by means of the following formulas:

\[
p_{pp} = p_{00} \quad (5)
\]

\[
p_{fp} = \frac{1}{N} \sum_{i=1}^{N} p_{i0} \quad (6)
\]

\[
p_{pf} = \sum_{i=1}^{N} p_{0i} = 1 - p_{00} = 1 - p_{pp} \quad (7)
\]

\[
p_{ff} = \frac{1}{N} \sum_{j=1}^{N} \sum_{i=1}^{N} p_{ij} \quad (8)
\]

2.5 Step 5 – stop criterion of the GA

Chosen stop criterion of the genetic algorithm is lack of improvement of found best solution.

3 Results

The methods have been verified using a low pass filter [2] (fig.2). Fault dictionary contains signatures of selected eight faults corresponding to faults of four elements (C1, C2, R2, R4) and a healthy circuit signature. Size of population is 30 individua. Length of the excitation vector is chosen as 100 and 200 samples. Required resolution of spectral analysis is the objective.

The method has been verified on three cases, which differ by CUT response feature extraction and spectrum optimisation of found excitation:
1) Euclidean classifier only
2) Euclidean classifier with spectrum optimisation
3) Euclidean classifier with wavelet analysis

3.1 Euclidean classifier only

Number of GA iterations: 113. Best solution is presented in the fig.4 and its spectrum in the fig.5. The spectrum has been averaged over whole signal length. Probabilities of fault location and detection for both unit step and obtained excitation are presented in tables 1-4.
3.2 Euclidean classifier with spectrum optimisation

Number of GA iterations: 61. Probabilities of fault location and detection for both unit step and obtained excitation are presented in Tables 5-8.
3.3 Euclidean classifier with wavelet analysis

Number of GA iterations: 40. Base wavelet for the classifier: Daubechies, 3rd order [1]. Probabilities of fault location and detection for both unit step and obtained excitation are presented in Tables 9-12.

### Table 9 Probability of fault classification for optimal excitation [%]

<table>
<thead>
<tr>
<th></th>
<th>D0</th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>D4</th>
<th>D5</th>
<th>D6</th>
<th>D7</th>
<th>D8</th>
</tr>
</thead>
<tbody>
<tr>
<td>F0</td>
<td>17</td>
<td>12</td>
<td>21</td>
<td>15</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F1</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>9</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F2</td>
<td>16</td>
<td>24</td>
<td>17</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F3</td>
<td>28</td>
<td>28</td>
<td>14</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F4</td>
<td>16</td>
<td>1</td>
<td>24</td>
<td>17</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>F5</td>
<td>16</td>
<td>1</td>
<td>24</td>
<td>17</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>F6</td>
<td>16</td>
<td>1</td>
<td>24</td>
<td>17</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>F7</td>
<td>16</td>
<td>1</td>
<td>24</td>
<td>17</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>F8</td>
<td>16</td>
<td>1</td>
<td>24</td>
<td>17</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

### Table 10 Probability of fault classification for unit step excitation [%]

<table>
<thead>
<tr>
<th></th>
<th>D0</th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>D4</th>
<th>D5</th>
<th>D6</th>
<th>D7</th>
<th>D8</th>
</tr>
</thead>
<tbody>
<tr>
<td>F0</td>
<td>16</td>
<td>11</td>
<td>35</td>
<td>24</td>
<td>17</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>F1</td>
<td>11</td>
<td>16</td>
<td>35</td>
<td>24</td>
<td>17</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>F2</td>
<td>11</td>
<td>16</td>
<td>35</td>
<td>24</td>
<td>17</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>F3</td>
<td>16</td>
<td>11</td>
<td>35</td>
<td>24</td>
<td>17</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>F4</td>
<td>16</td>
<td>11</td>
<td>35</td>
<td>24</td>
<td>17</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>F5</td>
<td>16</td>
<td>11</td>
<td>35</td>
<td>24</td>
<td>17</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>F6</td>
<td>16</td>
<td>11</td>
<td>35</td>
<td>24</td>
<td>17</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>F7</td>
<td>16</td>
<td>11</td>
<td>35</td>
<td>24</td>
<td>17</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>F8</td>
<td>16</td>
<td>11</td>
<td>35</td>
<td>24</td>
<td>17</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>9</td>
</tr>
</tbody>
</table>

By comparing results contained in tables, the following conclusions can be drawn.

Finding optimal shape of the excitation signal by means of Euclidean classifier without spectrum optimisation requires the smallest computational effort, significantly increases probability of correct fault location (tab. 1 and 2; main diagonals) and, in many cases, decreases risk of incorrect fault location (tab. 1; underscored entries outside main diagonal).

The same procedure with spectrum optimisation gives results similar to the ones found by previous method (both signal shape in time domain and spectrum). Elimination of high frequency components has not been achieved, which leads to conclusion that excitations with limited bandwidth have poor fault location and detection abilities. Such signals are missing in the final solution, because they have not “survived” the evolution process.

The procedure with additional use of wavelet analysis also returned excitation with shape and bandwidth similar to signals found by previous methods. Additionally, utilisation of wavelet transform caused interesting and desired effect of localisation of faults which are completely undetectable when wavelet transform is not applied (methods 1 and 2) (tab. 9 and 10; F2-D2 and F5-D5).
4 Conclusions

The presented methods have proven that fault detection and location can be significantly im-
proved by means of testing using specialised aperiodic signals. The results have been com-
pared with fault analysis using the simplest aperiodic signal: a unit step. It has been also ob-
served that wavelet classifier makes possible to locate (and detect) faults completely unde-
tectable by means of testing using unit step excitation. Importance of high frequency spectrum
components is another conclusion.

The presented approach has disadvantages too. Great computational complexity is the most
important. However, the methods belong to Simulation Before Test (SBT) class of testing
procedures, so computational time is not as crucial as in case of Simulation After Test (SAT)
procedure.

All above methods are in early development stage and require further work. Main directions
are reduction of computational effort and improvement of fault location and (especially) de-
tection. Nevertheless, the results obtained so far are promising.

Acknowledgments

This paper is supported by grant 3 T11B 03129.

References

[2] Golonek, T.: Utilisation of evolutionary computational techniques for designing and di-
agnosis analog electronic circuits (in Polish), PhD thesis, Silesian University of Tech-
nology, Gliwice, Poland 2003
Computational approaches to the electrical properties of the tin dioxide surface – effect of oxygen adsorption

Weronika Izydorczyk
Institute of Electronics, Silesian University of Technology,
Akademicka 16, 44-100 Gliwice
Weronika.Izydorczyk@polsl.pl

Abstract: An influence of oxidation of SnO$_2$ (110) surface on depleted layer raise has been studied by means of computer simulation. The total coverage and the coverages by the atomic and molecular oxygen ionic species versus temperature in dry synthetic air for different concentrations of single donors in the bulk were calculated. The surface potential and in-depth profiles of the electric potential and carrier concentrations in the depletion region, as well as the SnO$_2$ layer conductance, have been rigorously computed at various temperatures (from 400 to 900 K). The width of depletion layer $x_o$ as a function of single donor concentrations in the bulk were obtained.

1 Introduction

Tin dioxide is a semiconductor material widely used in various areas of technical applications due to its interesting optical, electronic and thermal properties. The n-type semiconducting behavior of SnO$_2$ layer arises from ionized donor impurities and oxygen vacancies in the bulk. It is well known in the case of metal oxide semiconductors, that oxygen in various forms plays a crucial role in gas sensing due to its reactivity. Surface of the sensor is then especially sensitive to reductive gases. The oxygen adsorbs on the surface in forms like O$_2$, O$_2^-$, O$^-$, O$^{2-}$. The investigations on SnO$_2$ surface using electron paramagnetic resonance (EPR) [1] indicates that below 200°C molecular species O$_2^-$ are present on the surface SnO$_2$, while atomic species O$^-$ become the dominating species above 150-200°C at low oxygen pressure. The ratio of O$_2^-$ to O$^-$ concentration increases with the oxygen pressure. Surface coverage by oxygen (surface concentrations per adsorption site) is independent on temperature and gas oxygen pressure for moderate value of oxygen pressure. [2]. According to investigations the oxygen ion O$^-$ plays special role between the catalytic surface processes and the electrical changes obtained in the temperature range from about 200°C to 500°C.

The oxygen adsorption is limited due to the band bending. The chemisorptions cannot take place anymore, if the Fermi level in the bulk is equal to the energy of the highest occupied surface states. Oxygen adsorption depends strongly on donor concentration in the bulk $N_d$ because increase of $N_d$ causes Fermi level in the bulk to bring closer to bottom of conductance band. The maximal coverage of the surface with oxygen results from the Weisz limitation (from $10^{12}$ to $10^{13}$ molecules per cm$^2$) [3].

The aim of this work is a rigorous theoretical analysis of the influence of surface oxidation on the conductance of SnO$_2$ layers with a thickness larger than the Debye length. From the one-dimensional numerical solution of the Poisson equation, the detailed in-depth profiles of the potential barrier $V(x)$ and carrier concentrations $n(x)$ were obtained for depleted layers. Then, the values of surface potential and layer conductance were determined. These magnitudes were calculated for different bulk doping and carrier mobility. Furthermore, the influence of temperature on near-surface region and conductance with assumption of non-mobile single bulk donors was studied. The results of the performed calculations are particularly important for better understanding of the mechanism of electronic processes in the near-surface region.
of the depleted SnO₂ surfaces used in gas sensors working at higher than room temperature conditions.

2 Computer procedure

The calculations of total coverage, density of trapped electrons on the tin dioxide surface as well as the width of depletion layer were carried out using a computer simulation for depletion layer based on the rate equations following Ref. [4]. These computations are based on the adsorption-desorption model proposed by Lantto et al. [4] where possible surface reactions and phenomena (adsorption, dissociation, recombination, desorption) are assumed to take place on the surface. The rate equations consider the electron transfer between different oxygen ionic species on the surface and the bulk conduction band. The negative charge trapped in oxygen species causes an upward band bending and thus, a reduced conductivity compared to flat band condition. Data concerning the surface coverage by oxygen ions O₂⁻ and O⁻ needed for simulations have been taken from Refs [4, 5].

The electronic surface and bulk parameters were obtained from the computer program, which realizes a one-dimensional numerical solution of the Poisson equation and gives the detailed in-depth profiles of the potential barrier V(x) and carrier concentration n(x) for different bulk donors concentration and temperatures. For strongly n-type materials such as SnO₂ the hole concentration may be neglected. The sample conductance is determined from classical formula [2, 6] using the obtained carrier in-depth profiles. The calculations were carried out for n-type SnO₂ with different doping level (non-mobile single bulk oxygen vacancies) from \( N_d = 3 \times 10^{14} \) cm\(^{-3} \) to \( N_d = 3 \times 10^{18} \) cm\(^{-3} \) and within the temperature range of 400K to 900K.

In calculations was assumed, that the bulk oxygen vacancies are singly ionized and non-mobile. The double ionization of oxygen vacancies giving rise to the mobile donors is observed above 1100°C for single crystals [7] or layers with grains with high surface to bulk ratio [8] (500°C and below).

The calculations of the surface energy barriers in the one-dimensional case for a planar semi-infinite surface are a good approximation for the surfaces of big grains (SnO₂ sensors), where a depth of the depletion layer is more smaller than thickness of the sample [9].

In the calculations, author assumed following parameters: electron mobility \( \mu_e = 150 \) cm\(^2\)/(Vs) at temperature 300K, effective mass of electron 0.3\( m_e \). The temperature variations in the carriers mobility and band gap were taken into account following Refs [10, 11].

3 Result and discussion

It is well known that SnO₂ surface oxidation leads to rise of depletion layer (potential barrier appears on the surface). The results of calculations of an influence of single donor concentration in the bulk and temperature on the total coverage of the oxygen ions of SnO₂ surface, in-depth profiles of the potential V(x) and surface barriers eV\(_S\), conductance per square and width of depletion layer are presented in Fig. 1-4. The calculated total coverage (surface concentrations per adsorption site) and the coverages of the oxygen ions O⁻ and O₂⁻ versus temperature in dry synthetic air, for single donors concentration \( N_d = 3 \times 10^{17} \) cm\(^{-3} \) are shown in Fig. 1a. The coverages of the atomic and molecular ionic species are equal at temperature about 445K. It should be noted, that the calculated value is comparable to transition temperature of 423K obtained from the measurements by Chang [12, 13] and by Rembeza et al. [14]. As resulting from Fig. 1 the total coverage depends on temperature and bulk donor concentrations according to paper [4]. The density of trapped electrons on the surface \( N_t = ([O^-]+[O_2^-]) \) increases from \( 9.5 \times 10^{10} \) cm\(^{-2} \) to \( 3.28 \times 10^{12} \) cm\(^{-2} \) with increasing donor concentrations from \( 10^{15} \) cm\(^{-3} \) to \( 10^{18} \) cm\(^{-3} \) in the bulk (Fig. 1b).
Fig. 1  a) Calculated coverage of the oxygen ions O¯ and O₂¯ and total coverage (O¯ + O₂¯) versus temperature in dry synthetic air, for single donor concentration Nd=3×10^{17} cm^{-3}, b) density of trapped electrons on the surface versus donor concentration in the bulk.

The detailed in-depth profiles of the potential V(x) and influence of donor concentration at the range of temperatures between 300K and 900K on eV_s are shown in Fig. 2. For instance the surface potential changes from −0.78V to −0.97V when temperature changes from 400K to 800K for constant donor concentration N_d=3×10^{16} cm^{-3}. Ionosorbed molecules that are associated with charge transfer from the solid to the adsorbed molecule will induce a band bending and consequently a variation in the conductivity. The surface barriers are shown to increase with increasing constant donor concentration at all temperatures. The surface barriers decreases (below 470K), then increases with increasing temperature for constant donor concentration. Furthermore, the calculated values of the surface barriers are consistent with those obtained by Lantto et al. [4].

Fig. 2 Depletion layer. In-depth profiles of the potential V(x) and surface barriers eV_s versus temperature for different donor concentrations in the bulk.
Contrary to the surface potential, the depletion region width is weakly sensitive to temperature but strongly to the doping concentration (Fig. 3b). In the range of changes of donor concentrations from \( N_d = 10^{15} \text{cm}^{-3} \) to \( N_d = 10^{18} \text{cm}^{-3} \), \( x_o \) decreases from value 1 µm to 30nm. On the other hand, the surface space charge region contributes significantly to the conductance of semiconducting layers. The sample conductance per square is determined from the classical formula [2] using the obtained carrier in-depth profiles (Fig 3a). The dependence of conductance per square on temperature is shown in Fig. 4. For samples with dimensions smaller than the space charge region and relative low carrier concentration a strong change in the conductance may be expected [15]. If thickness of SnO\(_2\) sample decreases, then ratio of thickness of depletion layer and thickness of the sample increases; for thin layers of SnO\(_2\) the sample can be depleted as a whole and resistance of the sample became large.
4 Conclusions

From computer calculations author revealed the strong influence of the depleted SnO$_2$ layers on the electrical parameters of the surface space-charge region, including the surface potential, density of the trapped electrons on the surface (surface charge), and width of the depletion layer. These results are important for the modeling of the electrical properties of SnO$_2$ thin film-based devices. The performed one-dimensional analysis is a good approximation for the surfaces of the large grain SnO$_2$ sensors, where the depletion layer width is much smaller than the grain diameter.

5 References

Response of SnO2 - based gas sensor structure to synthetic air and water vapour

Weronika Izydorczyk
Institute of Electronics, Silesian University of Technology,
Akademicka 16, 44-100 Gliwice
Weronika.Izydorczyk@polsl.pl

Abstract: The systematic experimental studies of the chemical composition and response to synthetic air exposure of the SnO2 based sensors were performed. The sensing SnO2 layers were deposited on alundum ceramics (Al2O3) using Rheotaxial Growth and Thermal Oxidation (RGTO) method. The in-depth chemical composition profiles were obtained using Ar+ ion sputtering and scanning Auger microprobe (PHI 600 system). The resistance measurements of SnO2 layers were made under the dry and humid synthetic air flow, during heating and cooling of the samples within the range of temperature from 20°C to 400°C. The obtained results showed that the sensor structures are highly sensitive to both oxygen adsorption and water vapour contain in air. On this basis, the relative changes of the sensor resistance versus temperature were determined. In particular, the sensor structure exhibited the maximum sensitivity under synthetic air exposure at the temperature about 275°C.

1 Introduction

Tin dioxide (SnO2) is a fundamental material for fabrication of the oxide-based gas sensors because of its interesting physical and electrical properties, like high sensitivity of a conductance to toxic gas adsorption and stability to heat treatment. Recently, our group developed a prototype of optimised SnO2–based sensor using Rheotaxial Growth and Thermal Oxidation (RGTO) method for detection of nitrogen dioxide, which is one of strongly toxic air pollution [1]. One important problem in successful mastering of RGTO techniques is the control of the preparation of sensory films, which should feature high sensitivity and selectivity, short response time and high response repeatability. The studies have shown that the sensor sensitivity depends mostly on the precise control of technological process parameters (cleanness, smoothness and temperature of the substrate during tin deposition, crystallite diameters, and active layer thickness) as well as on working conditions of the sensor (temperature, humidity and concentration of gas under examination) [1-3].

One important aspect of this work is the determination of factors (chemical composition, film structure and conditions of its preparation, operating temperature, concentration of the gas analysed) that have a decisive effect on the optimum operation of the sensor in the surrounding atmosphere, which first of all contains oxygen. Investigations described in the literature dealt mainly with the interaction of oxygen and synthetic air with thick films [4]. Not many investigators have studied the effect of chemical composition on sensory properties of thin films [5]. This paper shows that the films prepared are sensitive to oxygen contained in synthetic air and to the content of water vapour. Measurements were made of the resistance of SnO2 thin films in dry synthetic air during the heating and cooling of samples within the temperature range of 20 to 400°C. The effect of air humidity on sensor response was also studied. The obtained measurement results were used to calculate relative changes in resistance and the dependence thereof on sensor operating temperature was determined. The sensor showed maximum sensitivity to exposure to O2 when the operating temperature was ca. 275°C. The mechanism of gas molecule adsorption on the surface of the semiconductor is
in large part determined by the chemical and electron properties of the surface and near-surface region. Qualitative and quantitative analysis of the chemical composition of the SnO\(_2\) thin film was carried out by means of Auger Electron Spectroscopy (AES) in order to determine the effect of uniformity (compositional depth profiling) of the film on sensor response. It was found that uniform, well-oxidized films ensure repeatability of resistance measurement results under the conditions of interaction with synthetic air. Such films are also characterized by short regeneration time [6].

2 Experimental

The studied sensor structure was based on the SnO\(_2\) thin film grown in the RGTO process. The tin drops were deposited on alundum ceramics within the range of the substrate temperature from 250°C to 275°C. Subsequent process of thermal oxidation of Sn layers to the form of SnO\(_2\) was run in the atmosphere containing oxygen, at the temperature of 700°C. The resistance measurements of SnO\(_2\) layers were made under the dry (3% humidity) and humid synthetic air flow (50% humidity), during heating and cooling of the samples within the range of temperature from 20°C to 400°C. Resistance value was read after 5 minutes from the moment the desired temperature was reached, so that the system could attain equilibrium (resistance of the system was established). Dry and humid synthetic air was introduced to the testing chamber. The gas feeder created constant airflow inside the test chamber, which amounted to 500 ml/min. Precise control of the synthetic air mass flow was obtained by the use of Mass Flow Controller (MFC) supervised by microprocessor system. The measurements of sensor resistance were done using Agilent Multimeter 3497A. The temperature control was realised by means of Mastech M383 multimeter. The in-depth chemical composition profiles were obtained using Ar\(^+\) ion sputtering and scanning Auger microprobe (PHI 600 system). The ion energy was 3 keV and the sputter rate was about 7 nm/min.

3 Results and discussion

Fig. 1 compares the responses of a thin film structure obtained by the RGTO method with those of a thick film TGS 812 sensor in the process of heating and cooling in the temperature range.
Fig. 2  Relative change of resistance versus temperature a) in dry synthetic air b) in humid synthetic air

range of 20 to 300°C. Measurements indicate that during cooling the thin film sensor regains its initial resistance in a shorter time than the thick film sensor.

Fig. 2 shows the dependence of relative changes of resistance \( \frac{\Delta R}{R_0}(T) \) – where \( R_0 \) is the film resistance at temperature of 23°C) of the three structures on operating temperature. The graphs show that these changes are affected mainly by the operating temperature and by the temperature at which tin was deposited on the ceramic substrate during the RGTO process. Tin deposition temperature was 255°C in the case of sensor 1 and 270°C in the case of sensor 2, 3 (different thickness of the sensor layer).

The relative resistance changes reach a maximum in the temperature range of 260 to 285°C. Moreover, as indicated by Fig. 2, this maximum is shifted towards lower temperatures when air humidity increases.

Fig. 3  Oxygen chemisorption at grain boundaries of an n-type oxide semiconductor (e.g. \( \text{SnO}_2 \)); oxygen ions generate a depletion region and a Schottky barrier \( eV_S \)
In the literature, within the temperature range studied here, the maximum of relative resistance changes was observed at ca. 320°C in the case of thin films [7], and at lower temperatures (280°C) in the case of thick films [8, 9].

The change of the sensor response with temperature increase can be explained in terms of the chemical reaction course between the SnO₂ surface and adsorbate molecules. At the first stage of adsorption, a gas molecule undergoes physisorption (no exchange of charge with the substrate occurs), than O₂ undergoes ionosorption. The oxygen ions can form on surfaces of oxide semiconductors as a result of the transfer of electrons to the adsorbed oxygen (Fig. 3). The course of reactions within the examined temperature range is as follows [10]:

\[
\begin{align*}
O_2(gas) + e^- & \rightarrow O_2^-(ads) \\
O_2^-(ads) + e^- & \rightarrow 2O^-(ads) \\
2O^-(ads) + 2e^- & \rightarrow 2O^- (ads)
\end{align*}
\]

Chemical composition determination by means of AES may indicate, among other things, whether tin oxidation in the RGTO process was sufficiently long, as the repeatability of measurement results of sample resistance during gas adsorption depends on the complete oxidation of SnO₂ films [6]. Fig. 4 shows the Auger spectrum of the structure with tin deposition temperature of 255°C. Peaks representing oxygen O, tin Sn, carbon C, nitrogen N are apparent. The differentiated form of the spectrum was used to calculate oxygen to tin concentration ratio [O]/[Sn], which for the weakly oxidized sample varied from 1 (in the bulk of the sample) to 1.34 (on the surface). This indicates the presence of a mixture of tin oxide and tin dioxide. In the case of a well-oxidized sample this ratio reached the value of 1.83, indicating the prevalence of tin dioxide. Fig. 5, which shows in–depth profiles of the relative component concentration, indicates that the volume of the sample (up to the depth of ca. 140 nm) is uniform in terms of chemical composition.
4 Conclusions

AES examination may help select proper tin oxidation time (which depends on the thickness of tin film deposited by RGTO) to ensure that uniform, well-oxidized films are obtained, which in turn ensure repeatability of responses of the structures obtained to adsorbed gas. Depending on tin deposition temperature during rheotaxial growth, the responses of the structures obtained to the detected gas (synthetic air) may differ widely. The presence of water vapour in air causes a shift of the maximum of relative resistance changes towards lower temperatures in comparison to dry synthetic air. The data obtained enable optimization of sensor structure production with a view to attain high sensitivity, stability, repeatability and short time of response to the content of the gas detected in the environment.

Thorough studies, both experimental and theoretical, planned for the future, of sensory, electron and chemical properties of tin dioxide thin films (RGTO) will constitute an important stage in the understanding of the adsorption of oxidizing gases on the surface of SnO₂.

5 Acknowledgements

The author wish to thank Prof. B. Adamowicz, under whose guidance this work was performed, for the numerous and valuable discussions which were held on all the subjects involved in this study. I am grateful to Dr J. Uljanow for providing me with samples and to Dr W. Jakubik, Dr A. Klimasek for help to prepare experiments.

6 References


Laboratory thermo regulator for measuring the temperature
dependence of biomedical sensors

Pavel Kubina
Department of Measurement and Control, Faculty of Electrical Engineering
and Computer Science, VSB–Technical University of Ostrava
pavel.kubina@vsb.cz

Abstract: The article describes laboratory thermo regulator design. The thermo regulator is
intended for the measuring temperature dependence of biomedical sensors. The paper is
focused on one of the most important parameter in process control – temperature
measurement. Types of sensors and their comparisons are also mentioned. Interior
temperature of the thermo regulator is controlled by the thermoelectric Peltier’s module.
Temperature can be adjusted in range from 0°C up to 50°C.

1 Temperature regulation
Temperature is one of the most common measured quantity in control system. Accurate
measurement of the temperature is not so easy and to obtain accuracies better than 0.5 °C the
careful selection of temperature sensor is needed. Selection should be based on at least four
parameters: linearity, temperature range, sensitivity, and physical size. The temperature
sensor must provide an appropriate electrical signal to the controller. Due to the several
sources such as the sensor non-linearities, temperature gradients and calibration errors occur.

1.1 Temperature sensors
There are many types of sensors to measure the temperature An older classical sensors such
as the thermocouples, RTDs, and thermistors are extensively used due to their big advantages.
The new types of sensors as the integrated circuit sensors and radiation thermometry devices
are suitable only for short list of application.
Selecting the appropriate sensor is not always easy. This depends on many factors such as the
temperature range, required accuracy, speed of response, environment. Thermocouples are
best suited to very low and very high temperature measurements. They are low cost and vary
robust, external power is not required to operate and the typical accuracy is ±1°C. Thermocouples are usually used in high temperature industries. RTDs are used in medium
range temperatures, they offer high accuracy, typically ±0,2°C. Unlike thermistors, the
resistance of an RTD increases linearly with temperature. The operation of RTDs require
external power. RDTs are commonly used in lower temperature, higher precision chemical
industries. In medical applications are generally used thermistors. They are not as robust as
the thermocouples or the RTDs. The popularity of thermistors is mainly due to their
sensitivity and size. Thermistors change several hundreds of ohms per degree Celsius. Most
thermistor beads are the size of pin heads, allowing them to be precisely mounted to the
device being measured. Thermistors are low cost and their accuracy is around ±0,2°C.
Semiconductor sensors are used in low temperature applications. Their thermal coupling with
environment is not very good and the accuracy is around ±1°C. Some models offer digital
outputs enabling them to be directly connected to computer equipment without need of A/d
converters. Radiation thermometry devices measure the radiation emitted by hot object based
upon the emissivity of the object. But the emissivity is usually not known accurately and
additionally it may vary with time making accurate conversion of radiation to temperature difficult. The temperature range, advantages and disadvantages of various types of temperature sensors are shown in Table 1

Table 1  Comparison of temperature sensors

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Temp. range, °C</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>thermocouple</td>
<td>-270 to +2600</td>
<td>- wide operating temperature range</td>
<td>- non-linear</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- low cost</td>
<td>- low sensitivity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- rugged</td>
<td>- reference junction compensation required</td>
</tr>
<tr>
<td>RTD</td>
<td>-200 to +600</td>
<td>- linear</td>
<td>- slow response time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- wide operating temperature range</td>
<td>- expensive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- high stability</td>
<td>- current source required</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- sensitive to shock</td>
</tr>
<tr>
<td>thermistor</td>
<td>-50 to +200</td>
<td>- fast response time</td>
<td>- non linear</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- small size</td>
<td>- current source required</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- low cost</td>
<td>- limited operating range</td>
</tr>
<tr>
<td>integrated</td>
<td>-50 to +150</td>
<td>- some types can be directly connected to microprocessor</td>
<td>- self heating errors</td>
</tr>
<tr>
<td>circuit</td>
<td></td>
<td>- highly linear</td>
<td>- limited operating range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- low cost</td>
<td>- not good thermal coupling</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>with the environment</td>
</tr>
</tbody>
</table>

All these temperature sensors are used in biomedical and biophysical research application.

1.2 Laboratory thermo regulator design

The laboratory thermo regulator is intended above all for the measuring temperature dependence of biomedical sensors. Relatively small dimension of the laboratory apparatus are sufficient for this purpose. The high precision inside temperature control is required, thus the thermo electric cooler (TEC) is selected as a control member. Because of their solid-state construction, these small devices are very reliable and relatively easy to use. TEC control requires a reversible power source capable of providing positive and negative voltages. To accomplish this from a single supply, an H-bridge circuit is used.

The laboratory thermo regulator is constructed from a glued-laminated timber construction with cover. The outwalls are 12mm thick. The insulation quality is extremely important because significantly affects the cooling rate. Thus, the outwalls are widen with 32mm polyurethanes insulation board. Thermoelectric (TE) cooling system is mounted on the rear side of the lab case. The whole control electronic is placed on the cover of case.

1.3 Construction of the thermoelectric cooling system

The most critical aspect of system design when the components are chosen is the mechanical assembly. If the heat sink is not sufficient, or thermal conduction between components is poor, the system may not only be unstable but also has the potential to damage components.
The picture and detailed description of designed TE system is shown in Fig. 1. The system is composed of the cold and hot side. System is composed of cold and hot sides. Hermetically isolated Peltier module Al 011 \((Q_{\text{max}} = 50 \text{ W}, \Delta T = 85 ^{\circ} \text{C})\) is assembled between the sides of the system. Hot side consists of water copper block which exhaust the waste heat from the module. The waste heat is dissipated through the coolant in the reserator (designed by Zalman Tech Co., Ltd.) Cold side forms aluminium heatsink and fan. This heatsink absorbs the heat from internal case space in the cooling mode.

![Fig. 1: Configuration of TE system. 1 – cover strip, 2 – water CU block, 3 -Peltier module, 4 – spacer, 5 – heatsink, 6 – circulating fan](image)

Spacer is used to extend the distance between the hottest and coldest parts of the system by the maximum amount of insulation. The spacer is used on the cold side of the system due to the lower heat flux density. All of the mechanical interfaces between the loads to be cooled and ambient are thermal interfaces. All thermal interfaces tend to inhibit the flow of heat or add thermal resistance. Therefore all surfaces should have been burnished with a thin film of thermal grease. Excellent temperature stabilities can be achieved if the system is properly assembled. Compressive force is adjusted by screwing the middle set screw on the cover strip.

**1.4 Temperature monitoring and control**

Temperature control process consists of Peltier’s TE module and a temperature sensor which is used to sense the inner temperature and provide feedback to the microcontroller. The microcontroller runs a control algorithm and by way of MOSFET driver controls the TE module (see Fig. 2). The TE module is a DC device. Specified TE module performance is valid if a DC power supply is used. in practice, pulse wide modulation (PWM) is used widely (this method is particularly unsuitable for a TEC as a power source). The most common feedback configuration for high performance temperature controllers involves a combination of proportional, integral, and differential (PID) control. A PID type controller is implemented by the microcontroller PIC16F877 in order to achieve the desired output. The inner temperature is sensed using a thermistor. The sensor output is converted into digital form and is compared with stored desired temperature to form an error signal. A pulse width modulated waveform is generated from the microcomputer. The waveform is used to control a MOSFET driver MAX620. MAX620 driving an H-bridge switch that
controls the direction of passing current, thus whether the heat is absorbed or generated by the module.

1.5 Conclusion
The Peltier effect can be controlled linearly with electricity, the thermo electric cooler has been found in many applications involving precision temperature control. Described configuration of TE system is used in laboratory thermo regulator. The inner temperature can be adjusted in range from 0°C up to 50°C with 0,2°C accuracy. The construction solution notably effects performance of TE system. If the heat sink (or water block) gets too hot, efficiency of the Peltier device is greatly reduced. Liquid heat exchanger ensures an optional heat dissipation.

2 Acknowledgment

The support for this research work has been provided by the GAČR project 102/05/H525: The Postgraduate Study Rationalization at Faculty of Electrical Engineering and Computer science VSB-TU Ostrava.

3 References
[8] www.rmtltd.ru
Use of modular neural networks to fault diagnosis in analog active filters

Ireneusz Kurowski
Jerzy Rutkowski
Silesian University of Technology

Abstract: This paper presents a diagnostic system created for localization of faults in analog active filters. The system is based on dictionary method. As stimulation signals, dc supply voltage and sinusoidal input signal have been used. The frequency of sinusoidal signal is variable. Voltages and phases in output of the circuit are measured to create dictionary of faults. Radial Basis Function (RBF) neural networks were used as neural classifiers. Behavior Knowledge Space (BKS) was used to find the optimum final answer of parallel classifiers.

1 Introduction

Over last decades interest of analog fault diagnosis has been growing, because large electronic systems like mobile communications, process control or automotive systems usually implemented by digital techniques, often interface with the external world through analog devices. In digital implementations, well-consolidated techniques for automated test and fault diagnosis are used. Testing of analog circuits is complicated owing to the continuous nature of their input and output signals, tolerance and non-linearity of the circuit elements, presence of noise and the difficulty in modeling analog behavior. Testing of analog circuits is based on design engineer experience and on specification of circuits functionality. Test time is the critical element to reducing overall test costs, so the test system must approach the minimum test time limited by the device. Due to demand of a market, in the last years new methods of testing of analog circuits have appeared. They using new computational systems like a neural networks, fuzzy theory and evolutionary systems [1]. Nowadays, simulation before test is often used to create dictionary of faults. The computer simulations of “healthy” and “faulty” electronic circuit are realized before test. Next the dictionary of faults is created. The measurements, which are obtained while testing, are compared with dictionary of faults and next decision is made relative to appear of fault and its localization. This is very similar to the pattern recognition, where neural networks are used [2].

2 Fundamental proposition

The subject of this paper is diagnostic system created to localization of faults of analog active filters. Circuit before test simulations are performed at first, next dictionary of faults is created. Conditions that determine creation of a dictionary are presented below. As stimulation signals: supply dc voltage and input sinusoidal signal have been used. The frequency of sinusoidal signal is varying to cover whole band pass (stop) of tested analog active filters. Voltages and phases in output of the circuit are measured to create dictionary of faults. The typical faults were defined for all tested circuits, so both catastrophic and parametric faults were chosen. Simulations of electronic circuits were realized by SPICE. The obtained data were normalized and each state of circuit was binary coded to create dictionary of faults. In the same way testing file was generated. Then, dictionary of faults was
used to teach Radial Basis Function (RBF) neural classifiers [3]. Behavior Knowledge Space (BKS) was used to find the optimum final answer of the two parallel classifiers (parametric and catastrophic fault classifiers).

3 Teaching-testing files generator

The teaching-testing files prepared at the test stage have significant influence on neural classifiers. Computer software which was created to generate teaching-testing files contains information about structure and each state of circuit. Software generated file, is an input file for SPICE. This file is processed and output file is created by SPICE. Information contained in output file is used to create teaching-testing files. Teaching-testing files are generated in standard Matlab and they contain the following data: \( m \) - number of network inputs, \( n \) - number of network outputs, \( U \) - number of teaching vectors, \( D_{Uxm} \) - table of explanation data, \( T_{Uxn} \) - table of target data. The ranges of variation of test voltages and phases are very broad and diverse (fig. 1), so data were normalized. The following initial steps are performed:

- creating of two groups of data (catastrophic and parametrical faults),
- calibrating data in relation to non-faulty circuit,
- dividing of data into two groups (modulus and phases),
- calibrating data to the range of (0,1),
- binary coding each state of circuit.

Two normalized dictionaries of faults were received, first for catastrophic faults and second for parametric faults, so two neural classifiers were used to recognize a state of a circuit. The testing files were prepared in the same way.

4 Neural classifiers

Radial Basis Function (RBF) neural networks were used as neural classifiers. Simulations of neural classifiers were realized by Matlab and library Netlab, which was created by Ian Nabney [4] and Christopher Bishop [5]. Netlab is library functions and procedures, which enables simulations of neural networks and is available as „freeware” software [6]. The radial basis function network is the main practical alternative to the multi-layer perceptron for non-linear modeling [7]. The standard Netlab implementation of the RBF network includes one hidden layer with radial activation function and one output layer mostly with linear activation

![Fig. 1 The teaching file contains clusters of data of both catastrophic and parametrical faults](image-url)
function. The activation of the hidden units in an RBF network is given by a non-linear function of the distance between the input vector and a weight vector. Attraction of RBF network is that, there is a two-stage training procedure which is faster than the methods used to train multi-layer perceptron. In the first stage, parameters of the basis function are changed to model data density. The second stage of training determines weights in the output layer, and this is a quadratic optimization problem, which can be solved using methods from linear algebra. For RBF network, in certain circumstances, the network complexity can be matched to the data complexity to provide a simple way of determining the optimal size of the network. Netlab enable to create RBF network using function \texttt{rbf}, which has sequent parameters: number of inputs, number of hidden units, number of outputs and basis function, where the possible choices are ‘gaussian’, ‘tps’ (thin plate spline) and ‘r4logr’. The outputs (second-layer) of RBF networks in Netlab implementation are always linear. The returned network data structure \texttt{net} has the following fields:

- type always ‘rbf’
- nin number of inputs (\(m\))
- nhidden number of hidden units (\(H\))
- nout number of outputs (\(n\))
- nwts total number of weights and biases
- actfn string describing the hidden unit
- c hidden unit centres
- wi hidden unit widths
- w2 second-layer weight matrix
- b2 second-layer bias vector

To train the network, the function \texttt{rbftrain}, which containes the two-stage training algorithm, is called. To call this function, the user must pass the network \texttt{net}, input data \(D\) and target data \(T\) founded in teaching files as arguments. In the first stage the RBF network training algorithm is used to choose the basis function centres and, where appropriate, widths (for Gaussian basis functions). As the target is to model the data density with a linear combination of basis functions, an obvious approach is to treat the basis functions as a mixture model and to use the EM (expectation-maximization) algorithm to find the parameters. The second stage of the RBF network training algorithm is to compute the output weights, where a sum of squares error function was used. Since this error function is quadratic in the weights, its minimum can be found using the quadratic optimization methods, and the network is trained. Then is possible to compute the network outputs using testing files, so \texttt{rbffwd} function is used, which has sequent parameters: the taught network \texttt{net} and the testing data \(X\).

5 Behavior Knowledge Space (BKS)

Behavior Knowledge Space (BKS) was used to find the optimum final answer of parallel classifiers. BKS is the space which contains decisions for all used classifiers [8][9]. The space is \(K\)-dimension table, where each dimension is related to decisions of one classifier. Each classifier has \((M+1)\) possible decisions (\(M\)-quantity of patterns), which are chosen from file \{1, \ldots, M+1\}. The intersection of decisions of individual classifiers takes place of one cell of multidimensional BKS table. The cell which designates decision for the current sample is named central unit. Each central unit contains: quantity of all samples, quantity of classified samples for every patterns and the most representative pattern. The BKS space is defined by:

\[
\text{BKS}(e(1), \ldots, e(K))
\]

the BKS space, where classifier nr 1 gives decision \(e(1)\) and classifier nr \(K\) gives decision \(e(K)\),
The classification is passed with two steps, first is knowledge modeling, and second is making decision. During knowledge modeling the teaching file is used to design the cells of multidimensional BKS table, then values \( T_{e(1)\ldots e(K)} \) and \( R_{e(1)\ldots e(K)} \) for all cells are calculated. In the second step the central unit is assigned and next decision is made by the rule (1).

\[
E(x) = j \quad \text{when} \quad R_{e(1)\ldots e(K)}(j) = \max R_{e(1)\ldots e(K)}(m) \\
\text{and} \quad T_{e(1)\ldots e(K)} > 0 \\
\text{and} \quad \frac{n_{e(1)\ldots e(K)}(R_{e(1)\ldots e(K)})}{T_{e(1)\ldots e(K)}} \geq P_t
\]

(1)

where \( P_t \) from range \((0,1)\) is the trust threshold

\( E(x) = M+1 \) \quad \text{in the other case}

The knowledge modeling is executed only once to create BKS space. The classification decision is made always, when classified signal appears in the input of classifiers.

### 6 Test results

The studied circuit is band pass filter (fig.2) originally designed in [10]. The filter affirmed maximum 3 dB suppression for the band pass from 6,5 kHz to 15 kHz and minimum 20 dB suppression below 4 kHz and under 23 kHz. The filter was assembled of three operating amplifiers \( \mu A 741 \) and 15 RC components. Fault dictionary was generated for the circuit, and values of tolerance were postulated on a level of 1% for resistors and 2% for capacitors. The dictionary comprises both catastrophic and parametrical faults for all RC components. Output of circuit was chosen as a test node. First, the circuit was stimulated by the supply voltage, and the DC voltage in the chosen (output) node was measured, next sinusoidal signal was given to the input of the circuit, and voltages and phases at the output were measured. The frequency of sinusoidal signal was varying from 1 kHz to 26 kHz (26 test points). It enables to study performance within a whole band pass and stop of the tested analog active filter. That way number of measurements has been increased to 27 (26 AC + 1 DC) at the test node.

![Fig. 2 Band pass filter](image)
From SPICE simulations, measurement voltages and phases (53 different test data) were obtained, which were give at input of the neural network classifiers for all states of the circuit. Four kinds of faults were postulated for RC components. Their nominal values have been changed by –10%, +10%, –90% and +90%. That way classification of both deviation size and direction of changes are tested. The circuit is built of 15 RC components, so 60 faulty states and 1 healthy state have to be diagnosed. Each state of circuit was binary coded by 6-bits word, 2-bits to indicate size and direction of fault and 4-bits for localization of faulty component. The simulation was repeated 50 times for each of 61 state of the circuit. The teaching file, which has 3050 rows of data was created by computer simulation. Each row of the table contain 53 explanation data, which are input of neural classifier and 6 target data, which are expected on the output. In the similar way the testing file was created. The teaching and testing data were divided into two sections - catastrophic and parametrical, than were calibrated and normalized (see chapter 3). Two dictionaries of faults were received, so two RBF neural classifiers were used to classify catastrophic and parametrical faults. Each of them has 53 inputs, 6 outputs and one hidden layer with Gaussian basis function. Two group of teaching data were given on input of networks. At first, the EM (expectation-maximization) algorithm was used to choose the basis function centres and widths of a hidden layer. Than, the quadratic optimisation method was used to find the output layer weights. When the networks were taught, the knowledge modeling was executed to create BKS space (see chapter 5). During knowledge modeling the teaching file (containing both groups of data - catastrophic and parametric) was used to design the cells of two-dimensional BKS table, then values $T_{e(1)e(2)}$ and $R_{e(1)e(2)}$ for all cells were calculated. Next, the testing data were used to check results of classifications. The influence of a number of hidden layer units on the classification result has been studied. The data presented in fig. 3 were analyzed.

Fig. 3. The influence of a number of hidden layer units on the classification result
The used classifier has revealed some problems in parametrical faults localization – the average efficiency is near 80% for a number of hidden units more than 520 and less than 560. The average efficiency of localization of catastrophic faults (±90% of nominal) is more than 95% for a number of hidden units more than 520. The performance of classifier depends on position of RC components. The very important, from practical point of view is average percentage of correct detection of faults: 95,9% for parametrical faults and 99,8% for catastrophic faults.

7 Conclusions
The tests confirmed usability of the proposed neural system to localization of single catastrophic faults and to detection of single parametrical faults in analog active filters. The created software allows fast generation of teaching-testing files. The approach used DC and AC stimulation signals increases the quantity of teaching data, so it allows to decrease a number test nodes. It has been experienced that selection of basis function parameters, mean value and standard deviation, significantly influences fault classification. Two neural classifiers were used to recognize each state of circuit, first for catastrophic faults and secondary for parametrical faults. Behavior Knowledge Space (BKS) was used to find the optimum final answer of parallel classifiers.

8 References
Mathematical simulations of heat transfer in electrical sensors

Pavel Nevriva, Premysl Plesivcak
Department of Measurement and Control, VŠB-TU of Ostrava
premysl.plesivcak@vsb.cz

Abstract: In this paper, the heat transfer dynamics of the thermometer is discussed. The development of its mathematical model of sensor is presented, its numerical calculation and experimental verification of the mathematical model is introduced. In the real service, the sensor is screwed by its stem in an automobile oil pan case. It is used to control the temperature of oil. The stem of the thermometer is heated by oil while its opposite part with contacts is cooled by ambient air. To verify the mathematical model of the sensor, physical model of the actual system was built in laboratory. Mathematical model of the physical model was constructed and calculated numerically by the ANSYS environment. Reference measurement was made by infrared camera. Accuracy of the mathematical model was calculated.

1 Introduction

In this paper we have chosen bimetallic thermometer, used in oil pan case of the automobile. The bimetallic thermometer was selected as a typical case of a thermometer that is affected by the thermal energy. It serves as a pilot example for the larger project focused on the study of thermal load of thermometers. The paper deals with the mathematical simulation of temperature fields in the bimetallic thermometer, which is generated by the influence of the ambient temperature and oil in moving car. Mathematical model is created on the basis of laboratory physical model of the sensor. The thermal field is simulated in ANSYS environment. The plausibility of the mathematical model was tested by measurement.

2 Laboratory model of system

The mathematical simulation describes the simple physical model. It is rather difficult to accomplish the accurate mathematical model of thermometer build in actual engine in moving car. It is not simple to determine the ambient conditions of the sensor. The physical model of system is shown in Fig. 1. Oil pan case is replaced by thermostat with isolating wooden board. There is a natural convection between fluids and solids. The laboratory model was measured, than the convection coefficient $\alpha$ was corrected for the mathematical model. The bath and the ambient air are separated by the wooden desk with the orifice for the thermometer. The temperature of air is held on 294.2K. The temperature in oil bath is held on 365.2K.
The bimetallic thermometer is placed into three physical environments, see Fig. 1: air, wooden board, and oil. The housing of the bimetallic thermometer, see Fig. 2, is made from pressed brass MS60. Inside the hollow, there is the bimetallic strip with contact clamps. On the top of the sensor is the plastic cap with contact wires. In automobile the sensor is fixed to oil pan case, there is not a wooden board. The wooden board was used for simplicity. Parameters of the oil container are as follows:

- diameter of the container: 180mm
- high of the oil bath: 180mm
- thickness of the wooden board: 4mm

The thermometer is set to the initial temperature value near the ambient air temperature value and then plunged by its stem into the oil bath. Simultaneously, the temperature of the sensor is measured on the hexagonal contact part by infrared camera and the temperature is also controlled by set of temperature sensors. Measured data from the infrared camera are processes in MATLAB, accurately, in MATLAB image processing toolbox.

### 3 Mathematical model of the system

The analysis the heat transfer often combines three basic physical principles [3]: conduction, convection, and radiation. In this case the conduction suffices, because the temperature field is calculated in the solid body. Conduction could perform also in fluids, but it must be in inaction. The conduction describes the Fourier equation (1), that include initial and boundary conditions.

\[
\frac{\partial T}{\partial t} = \frac{\lambda}{\epsilon_p \cdot \rho} \left( \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} \right)
\]

where \( \lambda \) is thermal conductivity [Wm\(^{-1}\)K\(^{-1}\)], \( \rho \) is density [kgm\(^{-3}\)], \( c_p \) is specific heat capacity [J.kg\(^{-1}\).K\(^{-1}\)], \( T \) is temperature [K]. The initial condition is a temperature in the beginning of the process (the temperature at the time \( t = 0 \), \( T(0) = 294.2K \) (21°C)).

Boundary conditions are described by following formula of density of heat flow rate

\[
q = \alpha(T_A-T_S)
\]

where \( q \) is density of heat flow rate [Wm\(^{-2}\)], \( \alpha \) is heat transfer coefficient [Wm\(^{-2}\)K\(^{-1}\)], \( T_A \) is ambient temperature [K], \( T_S \) is temperature of solid surface [K]. Bimetallic thermometer and the board are affected by two fluids. Both the density of heat flow rate are calculated using the formula (2). The temperature \( T_A \) and \( T_S \) are measured. For the calculation of heat transfer coefficient \( \alpha \) the similarity method is applied. Calculated and measured values of the boundary conditions are written below in Table 1. Any parameters can be seen also in Fig. 1.
Table 1 Parameters of boundary conditions

<table>
<thead>
<tr>
<th>Boundary of the solid</th>
<th>$T_A$ [K] (°C)</th>
<th>$\alpha$ [Wm$^{-2}$K$^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boundary air-brass/plastic</td>
<td>294.2 (21.0)</td>
<td>$\alpha_{\text{air}} = 18.7$</td>
</tr>
<tr>
<td>Boundary oil-brass</td>
<td>365.2 (92.0)</td>
<td>$\alpha_{\text{oil}} = 83.5$</td>
</tr>
<tr>
<td>Boundary air-board</td>
<td>294.2 (21.0)</td>
<td>$\alpha_1 = 6.6$</td>
</tr>
<tr>
<td>Boundary oil-board</td>
<td>365.2 (92.0)</td>
<td>$\alpha_2 = 76.9$</td>
</tr>
</tbody>
</table>

4 Simulation and measurement of the system

The simulation of temperature field of the bimetallic thermometer was modeled in ANSYS environment, which is based on the finite element method. The finite element method is numerical method. ANSYS solution also provides the CFD analysis (Couple Field Analysis). The geometrical drawing of the sensor was made directly in ANSYS, see Fig. 3a. The wooden board is also simulated, but isn’t depicted here. The CFD analysis is used, because sensor is filled up of air inside. Also the transient analysis is applied. For each component material constants (thermal conductivity, specific heat capacity, and density) were specified. Then, finite element type (FLUID 142) was defined. After that, the geometrical model was meshed with previous attributes, see Fig. 3b. Finally, we apply the initial and boundary conditions, and solve the model.

For the measurement, two bimetallic thermometers were selected. These sensors were of the same run. Measurement was made by calibrated infrared camera and contact thermometers. The infrared camera was set approximately 20cm from the thermometer. The experimental configuration is shown in Fig. 4. The representative results can be read in the camera image only from the face plane of the hexagonal thermometer's contact part. The face plane was set to be perpendicular to optical axis of camera. Surface of the brass and the plastic was blacken to have the maximal emissivity coefficient of the surface $\varepsilon = 0.96$. Parameters of the infrared camera are: Sampling frequency: two frames/s, Temperature ranges: 230-390K; 270 - 770K; 620-1770K, Temperature sensitivity: 0.08K, Image resolution: 240x320 pixels, Spectral range: 7.5-13 µm, Accuracy of non-calibrated camera: 1K.
5 The comparison of simulated and measured data

Following figures represent the measured and simulated temperature fields of the contact part of the sensor in selected time, see Fig. 5. In data generated by infrared camera, the accurate values of temperature can be read only from the face plane of the hexagonal element.

![Fig. 5 Temperature field in 650s – (a) calculated, (b) simulated](image)

The measurement was repeated 12 times. The maximal differences between steady-state values of temperatures measured in the face plane at the hexagonal contact parts of thermometers were less then 0.4K. Fig. 6 shows the transient responses calculated and measured in the centre of the face plane. The total difference between calculated and measured values in temperature in steady state is approximately 1.5K, it is about 2.5%. The total difference in dominant time constant is approximately 20s, it is about 5%.

![Fig. 6 The temperature transient response](image)

6 Conclusion

In this paper, the simulation of the temperature dynamics of the bimetallic thermometer was presented. The task was described by the Fourier partial differential equation. The simulation was made in ANSYS. Results of the simulation were verified by the measurement made by the infrared camera. The results obtained lead to the conclusion that data from well conducted experiment have better accuracy than data generated from mathematical model. On the other hand, mathematical model can find temperature fields of products that are still in the design stage, generates temperatures inside the solid body, and has many other advantages. *The work and the contribution were supported by the project from Grant Agency of Czech Republic - 102/06/0498, Simulation of Modern Sensors' Electronic Thermal Load.*

7 References

Design of complex digital circuits in aspect of disclosure emission with use of simplified mathematic emanation model

Marcin Szelest
Evatronix S.A/Silesian University of Technology
szelest@best.gliwice.pl

Abstract: The use of FPGA and ASICs in several cryptographic application has many open issues which should be taken into consideration. This document includes some information about improving design flow for disclosure emission analysis. The analysis gave answer if developed device met security guidelines or not. This analysis can be made with simplified mathematic emanation model proposed in this article.

1 Introduction

Common digital circuits designed these days (ASIC, FPGA, SoC or typical microcontroller devices) often realize confidential functions or functions protected by patent law. In the first impression people think, that miniaturization in silicon industry gave sufficient safety for implemented algorithms, but nothing more erroneous[2][3][6]. Very important thing is way of designing circuits, which met defined specification, and additionally generate as less disclosure emission as possible. Concept of disclosure emission defines all possible influences with environment, which gave information about changes that occur inside observed device. In case of silicon circuits, this emission is generated by current’s impulses connected with CMOS gates switching. Each change of logical level causes generation of current impulse and associated with it electromagnetic disturbance.

To observing disclosure emission connected with current impulse cryptanalysis uses two basic methods: SPA (Simple Power Analysis)[7] and DPA (Differential Power Analysis)[5][7], however electromagnetic disturbance is analyzed by SEMA (Simple Electromagnetic Analysis)[4] and DEMA (Differential Electromagnetic Analysis)[1].

2 Design Flow Overview

Design flow for digital devices consists of consecutive stages, in which designer moves toward more complex description of designed circuit. As this description is more detailed, the abstraction level is lower.

Disclosure analysis bases on mathematic model of radiating silicon device. The most important thing for this analysis is information about changes of logical signals inside IC structure, so it entails necessity of pointing abstraction level most suitable for disclosure analysis. Table 1 shows short summary of abstraction levels[8] with opinion about them usefulness for the disclosure emission.

<table>
<thead>
<tr>
<th>Abstraction level</th>
<th>Short description</th>
<th>Usefulness for analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>System Level</td>
<td>It is the highest possible</td>
<td>For the sake of very low details</td>
</tr>
</tbody>
</table>
abstraction level, all descriptions are made in natural human language. Only global functionality is defined without any complex details.

**Algorithmic (Transaction Level Modeling - TLM)**

TLM is description method which focus on details of communication between functional modules. This way of modeling separates functionality of modules from functionality of communication between them. This level makes possible as well simulation as well observation of internal states, but it focuses on communication details. It causes that disturbance analysis would be reduced only to events on buses and external interfaces. In fact, on this level of abstraction, designer can only implement specialized communication protocols which gave high level of protection.

**Register Transfer Level (RTL)**

RTL level bases on registers, conditional statements and logical operations. Currently it is the lowest level of abstraction, which can be easily understood by human. Further levels are usually generated from this description with use of specialized computer software in process called logical synthesis.

**Logic Level**

At this level whole circuit is defined as connections between logical gates and other logical primitives. Functionality of modules is completely not distinguished from communication between them.

and lack of simulation possibilities this abstraction level has very slight usefulness for disclosure analysis. Only one thing can be taken into consideration - selection of specialized algorithms, which ensure small diversity of internal states in developed device. This level makes possible as well simulation as well observation of internal states, but it focuses on communication details. It causes that disturbance analysis would be reduced only to events on buses and external interfaces. In fact, on this level of abstraction, designer can only implement specialized communication protocols which gave high level of protection.

This level is much more detailed then TLM but still to complex to analyze. The main drawback is fact, that different synthesis tool will produce different netlists. Despite of this, at this abstraction level designer can define some HDL coding rules which make design a bit safer. This rules bases on symmetry of conditional instructions, which can be checked during compilation process. In case when symmetry is not fulfilled appropriate information should be displayed as warning message.

This is the last common level for ASIC and FPGA, and the most suitable level for disclosure analysis. The description is defined as pure netlist or as netlist with SDF timing file (after place&route), so all information about signal changes are easily accessible from testbench environment during postsynth simulation. Additional advantage is possibility of using testbenches form RTL level design with very low effort needed for
modifications. Choice of this level is compromise between effective simulation and satisfactory accuracy. Detailed description makes possible to obtain very accurate simulation, but at these abstraction levels, amount of details makes impossible efficient simulation.

### Circuit Level
At this level whole circuit is defined as connections between transistors and other semiconductive devices. Functionality of modules is completely not distinguished from communication between them. Detailed description makes possible to obtain very accurate simulation, but at these abstraction levels, amount of details makes impossible efficient simulation.

### Silicon Level
This is the lowest possible abstraction level. All transistors and connection between them are defined as geometric figures placed on silicon structure. The same features as in Circuit Level

### 3 Features of silicon device emanation
Digital circuit, build with use of CMOS technology, emits energy only during switching state of logical gates. This energy is emitted, in very wide spectrum of frequency (from radio waves to infrared emanation), as result of capacitance reload in pair of unipolar transistors. The spectrum characteristic depends on gates output load, length of connection between gates and complex relationship between capacitive and inductive coupling.

Generally disclosure emission in the close neighborhood of working silicon device can be defined as superposition of all transitions in specified period of time. Note that due of complexity of the common digital circuits real and accurate simulation is almost impossible. Even in case of simulating real physical phenomena on the layout level, designer is not able to predict what will happen if somebody for example draw the wire under IC package or put IC package in such place on PCB, where electromagnetic field is very strong.

### 4 Assumptions for mathematic model of disclosure emission
Due the impossibility of accurate simulation, some assumptions for mathematic model must be taken. Firstly assume that for disclosure emission analysis, sufficient information can be only knowledge if examined circuit is vulnerable for SPA/EMA attacks or not. Moving form accuracy simulation toward susceptibility-based analysis significantly reduces simulation effort and mathematic model accuracy.

As mentioned before the most suitable abstraction level for the disclosure emission analysis is the Gate Level threatened as netlist with post-simulation timing model. The netlist can be easily simulated with use of almost the same testbench as on RTL level. On the first step every signal in the netlist is attributed with defined value. This value should reflect to gate load, so it is dependent from number of gates input connected with current signal and average length of this connection with some random value derived independently, which models others not possible to predict relationships after place&route operation. Note that this random value is constant for one signal, only for time of one simulation. There is no possibility that two
different simulations will give the same results. Of course different simulation results does not disturb to determine if examined device is disclosure emission proof or not.

Figure 1 shows example structure of small part of the netlist. Gates A and D drives only one input (A drives D and D drives E), connection A-D is internal while connection D-E is external. Because of possibility that external connection will be routed with longer path that internal connection, value assigned for this net should be bigger then for the second. Connections A-D and E-H looks similar but they will be assigned with different values because of randomization. The randomization causes that results of \( n \) simulation on one specified design newer will be the same.

After attributing values to the signals from the netlist, the whole design is simulated, in post-synth mode with timing file (SDF)[9], with use of chosen simulation tool (Modelsim, NCSim or similar). All transitions of signals from the netlist are registered in waveform file with information about time when specified change happens. After whole simulation the waveform file is read again, and according to it's contents and values of attributes, analog waveform is generated. As more flat obtained waveform is as examined circuit is more disclosure emission proof.

5 Design flow with taking disclosure emission into consideration

Proposed design flow, which takes into consideration disclosure analysis, is shown on Figure 2.
Figure 2 Proposed Design Flow

Note that typical design flow is expanded by extra feedback loop, which is used to checking if device meets specified security rules. Modifications can be made both on RTL (handmade) or on the netlist (automatic – human independently) level. Some basic methodology for adding extra logic is described in [10]: SABL (Sense Amplifier Based Logic), SSDL (Simple Dynamic Differential Logic) and DWDDL (Divided Wave Dynamic Differential Logic).

During simulation appropriate test sequence should be forced. For microprocessor device it should be software, which realizes default secure operations. For hardware devices it should be collection of patterns which makes that simulated device works in default conditions.

6 Conclusions

The article shows survey of difficulty connected with building secure DPA and EMA resistant circuits. Special attention was taken to select appropriate abstraction level, which enables analysis of complex design both in ASIC and FPGA technology. Additionally, modification of typical design flow was proposed. The modification, thanks to pseudo-random simulations, enables automatic indication of parts of the structural description (netlist), which are especially responsible for undesirable disclosure emission. Thanks to operating on the most suitable abstraction level and because of use simplified mathematic model, disclosure analysis take only a bit more time effort, then post place&route simulation.
Proposed methodology makes possible to simulate and development very complex circuits (such as Control Processing Units form microprocessors). A very important feature is, that almost all details of implementation and netlist modifications can be hidden from engineer. From the designer point of view, only one level between logical synthesis and implementation is added.

7 References

Section 3

Information Technology
Adaptive Thresholding in Fuzzy Approach to Segmentation of Cruciate Ligaments

Paweł Badura
Silesian University of Technology
pawel.badura@polsl.pl

Abstract: This paper presents an approach to segmentation of the cruciate ligaments of the knee joint. Segmentation of such ligaments is difficult due to poor visibility of edges in some cases of injuries and their appearance on a small number of Magnetic Resonance Imaging (MRI) slides. In the current study the method based on fuzzy connectedness principle has been applied. It assigns a strength of connectedness every to possible pair of image elements and creates the connectivity scene. Then the scene is thresholded to produce final segmentation result. In order to minimize the user interaction needed an adaptive threshold selection method has been implemented and tested.

1 Introduction

The cruciate ligaments are the main stabilizers of the knee joint – the most complex joint in the human organism. The anterior (ACL) and posterior (PCL) cruciate ligament connect the femur with the tibia. They form a cross and their names refer to their tibial attachments [6]. Their susceptibility to injuries, occurring in both athletes and nonathletes, make a proper diagnostics a very important matter. The ligaments are visible at the Magnetic Resonance Imaging (MRI). In current research the sagittal T1-weighted images are studied. Segmentation of the ligaments is a difficult task due to poor visibility of edges in some cases of injuries. Other significant problem is their appearance on a small number of slides – in case of a 4mm distance between slides the PCL appears in 2-4 slides, yet the ACL is located on 1-2 slices. The ligaments can be seen as oval, narrow, dark structures, attached to the femoral and tibial bones [3]. Healthy PCL is almost black and slightly turns in the femoral part. The ACL has a bit higher intensity than the PCL and forms a straight structure. Injured ligament becomes brighter, wider, and less distinct from other tissues of the knee joint [4]. There are also some secondary signs of a ligament tear, like changes in its shape and direction or shifts, bruises and fractures of the bones.

The segmentation process is based on fuzzy connectedness principles [9,11]. This approach has been used in several image segmentation issues in both, MRI and Computed Tomography (CT) [9]. In this method the fuzzy connectivity scene is created based on fuzzy relations between image elements. Then, the scene is thresholded to produce final segmentation result. The aim is to minimize the user interaction needed to properly delineate the object. In order to achieve it an adaptive threshold selection method has been implemented and tested. It makes the selection of seed points the only interaction needed. Fuzzy connectedness, relative fuzzy connectedness and adaptive threshold selection method will be described in section 2. Section 3 presents the details of the implemented algorithm, along with some visual 2D and 3D examples of its application to the ligaments segmentation. Finally, in section 4 concluding remarks and propositions for the method’s improvement are pointed out.
2 Fuzzy connectedness and adaptive threshold selection

2.1 Fuzzy connectedness

The idea of fuzzy connectedness has been proposed in [11]. In this idea the fuzzy relation \( \rho \) between every two spels (spatial elements, pixels or voxels) in the image is defined. Let’s assume, that for a set \( C \) (image) and any two its elements \( c \) and \( d \):

\[
\rho = \{(c,d), \mu_\rho(c,d)\}|(c,d) \in C \times C\tag{1}
\]

\( \mu_\rho \) is a fuzzy membership function, so \( \mu_\rho \in [0,1] \). The relation \( \rho \) has to be reflexive

\[
\mu_\rho(c,c) = 1, \forall c \in C\tag{2}
\]

and symmetric

\[
\mu_\rho(c,d) = \mu_\rho(d,c), \forall (c,d) \in C \times C\tag{3}
\]

The basic relation presented in [11] is called fuzzy spel affinity \( \kappa \). For every two elements a value of \( \mu_\kappa(c,d) \) is assigned. It is based on their coordinate adjacency, intensities, gradient, and perhaps even their locations within an image. The general form of \( \mu_\kappa(c,d) \) is as follows:

\[
\mu_\kappa(c,d) = \mu_\alpha(c,d) \cdot [\mu_\psi(c,d), \mu_\phi(c,d), c,d]\tag{4}
\]

where \( \mu_\alpha \) is an adjacency relation (like hard 2D 4-neighbourhood or 3D 6-neighbourhood), \( \mu_\psi \) represents the intensity-based function, and \( \mu_\phi \) represents the intensity gradient-based part of the affinity. Several possibilities for (4) have been shown and described in [8]. The most often used form of \( \mu_\kappa \) is:

\[
\mu_\kappa(c,d) = \mu_\alpha(c,d) \cdot [\omega_1 \cdot h_1(f(c),f(d)) + \omega_2 \cdot h_2(f(c),f(d))]\tag{5}
\]

with weights \( \omega_1 + \omega_2 = 1 \), and \( h_1, h_2 \) being some membership functions corresponding to \( \mu_\psi \) and \( \mu_\phi \), respectively.

Fuzzy affinity as described above is nonzero only for the adjacent spels. We call any pair of adjacent spels \( c, d \) a link, and the value of \( \mu_\kappa(c,d) \) – its strength. Any sequence of spels \( <e_1, e_2, ..., e_m> \) such that for any \( i \in [1..m-1] \) a pair \( <e_i, e_{i+1}> \) is a link, we call a path. It is noted \( p_{cd} \) if \( c=e_1 \) and \( d=e_m \). The strength of a path is the strength of its weakest link – the smallest affinity along the path:

\[
\mu_\kappa(p_{cd}) = \min[\mu_\kappa(c_i,c_{i+1})]\tag{6}
\]

All paths \( p_{cd} \) between any two image spels \( c \) and \( d \) form a set \( P_{cd} \). The fuzzy connectedness is a fuzzy relation \( K \) between any two image spels \( c \) and \( d \) with membership function \( \mu_\kappa(c,d) \) being the strength of the strongest path \( p_{cd} \in P_{cd} \):

\[
\mu_\kappa(c,d) = \max_{p_{cd} \in P_{cd}}[\mu_\kappa(p_{cd})]\tag{7}
\]

The fuzzy \( \kappa\theta \) object with affinity \( \kappa \) and some threshold \( \theta \in [0,1] \) is defined as a set \( O_{\kappa\theta} \) of spels such that for every \( c, d \in O_{\kappa\theta} \) \( \mu_\kappa(c,d) \geq \theta \). It has been proven, that classification of a fuzzy \( \kappa\theta \) object does not require the computation of \( \mu_\kappa \) for all possible pairs of spels in \( C \). First some seed spel \( o \) is chosen, which is supposed to belong to an object \( O_{\kappa\theta(o)} \). Then, the connectivity scene \( C_o \) is computed for all \( c \in C \):

\[
C_o(c) = \mu_\kappa(o,c)\tag{8}
\]

with \( C_o(o)=\mu_\kappa(o,o)=1 \) according to (2). For each \( c \), \( C_o(c) \) describes the degree of its connectedness to \( o \). Thresholding the connectivity scene at some \( \theta \) gives a fuzzy \( \kappa\theta \) object \( O_{\kappa\theta(o)} \). A binary object \( O_{\kappa\theta(o)} \) – result of the segmentation process – is defined as follows:
For every two spels \( c, d \in O_{\kappa(\theta)} \) \( \mu_K(c,d) \geq \theta \), and for each spel \( e \not\in O_{\kappa(\theta)} \), \( \exists c \in O_{\kappa(\theta)} \) such that \( \mu_K(c,e) < \theta \). Moreover, if \( c \in O_{\kappa(\theta)} \), then \( O_{\kappa(\theta)} = O_{\kappa(\theta)} \). It means, that for fixed fuzzy affinity relation and threshold it does not matter which point from an object is taken as its seed. If a set \( O \) of \( M \) seed spels \( o_i \) is indicated rather than a single point \( o \), then the fuzzy connectivity scene \( C_o \) for such set is defined as the union of the connectivity scenes for all \( o_i \):

\[
C_o(c) = \bigcup_{o_i \in O} C_{o_i}(c) = \max\{\mu_K(o_i, c)\}
\]

2.2 Relative fuzzy connectedness

In [7,10] the idea of relative fuzzy connectedness has been introduced. The new object is indicated by marking its seed point \( b \) (or a set \( B \) of points \( b_i \)) outside the \( O_{\kappa(\theta)} \). New object \( O_{\kappa(\theta)} \) is treated as a background. Two connectivity scenes are then determined for both, the object \( Co \) and the background \( C_b \). Finally, each spel \( c \) is examined, whether his connectedness to \( o \) is greater than to \( b \). If so, it is included into \( O_{\kappa(\theta)} \):

\[
O_{\kappa(\theta)}(c) = \begin{cases} 1 & \text{if } C_o(c) > C_b(c) \\ 0 & \text{otherwise} \end{cases}
\]

(11)

It is possible to define more than two growing objects and let them compete for spels [7].

This method does not require a threshold selection, but in segmentation of difficult structures like cruciate ligaments it has also few disadvantages. It is sensitive to the seed points selection and fuzzy affinity determination and requires indication of larger number of such points. Moreover, it takes twice as much time as in conventional fuzzy connectedness. So, an easier method has been implemented here. Instead of determining the background fuzzy connectivity scene \( C_b \), the object fuzzy connectivity scene \( C_o \) is computed until the point \( b \) (or \( b_i \in B \) with the highest \( C_o(b_i) \)) is reached. There is no need to compute the full \( C_o \), because for all spels \( c \), which has not been analyzed before \( b \):

\[
C_o(c) < C_o(b)
\]

(12)

and so:

\[
C_o(c) \leq C_b(c)
\]

(13)

which fulfills (11), that \( c \) will be numbered among the background. A resulting fuzzy connectivity scene \( C_o \) is nonzero only for spels \( c \) with fuzzy connectedness to \( o \) higher than between \( o \) and \( b \). It does not mean, that all those spels should be treated as an object. The threshold for \( C_o \) has to be defined to determine and delineate an object.

2.3 Adaptive threshold selection

The idea of adaptive threshold selection for fuzzy connectedness method has been presented in [2]. It is based on entropy of the fuzzy connectivity scene. The entropy of an image \( X \) of size \( M \times N \) and \( L \) histogram levels of intensity is defined as:

\[
H(X) = \frac{1}{MN \ln 2} \sum_m \sum_n E(\mu_X(x_{mn}))
\]

(14)

where

\[
E(\mu_X(x_{mn})) = -\mu_X(x_{mn}) \ln(\mu_X(x_{mn})) - (1 - \mu_X(x_{mn})) \ln(1 - \mu_X(x_{mn}))
\]

\[
x_{mn} \in X; m = 1,2,...,M; n = 1,2,...,N.
\]

(15)

The following fuzzification method is employed:
where $t$ is a current threshold value, $h(l)$ is a histogram with $L$ levels and $K$ is a constant value such that $0.5 \leq \mu_X(x_{mn}) \leq 1$. The level $l_i$ for which the entropy $H(X)$ reaches its minimum is selected as a threshold level. Such an algorithm is used to select the threshold for the fuzzy connectivity scene $C_o$ computed as in 2.1.

### 3 Algorithm and results

The sagittal T1-weighted MR image is treated as a 3D matrix $C$. It is normalized to a range $[0,1]$. Optionally the region of interest (ROI) can be marked. Then a seed slice is chosen and seed points are selected on it – $o$ belonging to an object and $b$ in a background. Sets $O$ and $B$ can be chosen as well. With such a data a modified Dijkstra algorithm [1,5] has been implemented. It is listed below:

Input data: a 3D image intensities in matrix $C$, a fuzzy affinity $\kappa$, sets $O$ and $B$ of seed spels. Output data: a fuzzy connectivity scene $C_O$ with zero values for obvious background spels. Auxiliary data structure: a set $Q$ of spels $c \in C$.

1. Insert all $o_i \in O$ to $Q$.
2. Set $C_O$ to zero except spels $o_i \in O$, for which set $C_O(o_i)=1$.
3. Remove $d$ if $C_O(d)$ reaches maximum in $Q$.
4. If ($d \notin B$):
   5. For each spel $c$ adjacent to $d$:
      6. $v = \min\{C_O(d), \mu_X(c,d)\}$.
      7. If $v > C_O(c)$:
         8. Put $c$ into $Q$.
         9. Set $C_O(c)=v$.
   10. If $Q$ is empty: stop,
       Else: go to step 3.
11. Else:
   12. For each spel $c$, for which $C_O(c) \leq C_O(d)$ set $C_O(c)=0$ and stop.

Then an adaptive thresholding part of an algorithms is implemented according to (14-16). After the morphological closing operation (on a binary image $O_{\kappa \theta(o)}$) of segmented ligament its borders are plotted on original image. It is advisable to leave the possibility to manually adjust the threshold level on the basis of visual evaluation of the segmentation result.
An example of the performance is shown on figures 1 and 2. Fig. 1 shows the anterior cruciate ligament segmentation process of the 2D MRI slice of the knee joint. Two points inside an object (+) and two points in the background (×) are selected (fig. 1a). Fig. 1b is the visualization of connectivity scene $C_O$, $C_O$ and its histogram (for 100 intensity levels) are mapped into range [0,1]. Entropy has its minimum for $t=0.66$ and a segmentation result for such value is shown on fig. 1c. The 3D posterior cruciate ligament segmentation process is presented on fig. 2.
4 Concluding remarks

The paper describes a method of adaptive threshold selection in the fuzzy connectedness approach to segmentation of the cruciate ligaments. It has been tested on a set of 20 images with satisfactory results, but also following conclusions, requiring further research. The method is sensitive to seed points selection – the more points are selected, the better the result is. A modified automatic method for fuzzy affinity definition has to be proposed. Furthermore, a 3D image interpolation algorithm seems to be indispensable due to different voxel size in XYZ directions. Finally, the knowledge about the anatomy of the ligaments and their appearance can be used, e.g. in preprocessing or a fuzzy connectivity scene computation. The goal is a specified tool for computer aided diagnostics of the cruciate ligaments at MRI, enabling segmentation and visualization of healthy or injured ligaments.

5 References

The Application of Support Vector Machines and Kernel Methods in Speaker Recognition

Mariusz Bąk
Silesian University of Technology
Mariusz.Bak@polsl.pl

Abstract: The paper presents the idea of Support Vector Machine. SVM is a system for learning binary linear classifiers, which was introduced by Vapnik in 1963. The essential advantage of SVM classifier is that the resulting partition of the input space is optimal in terms of minimizing the structural risk and that the obtained solution is sparse and computationally efficient.

This paper contains a short description of SVM foundations and enumerates its main advantages and disadvantages. Because a concept of SVM is closely related to the use of kernel methods, a short overview of these methods is also included.

1 Speech as a biometric

Identification (i.e. determining the identity of a person) and verification (i.e. verifying whether an identity of a person is in accordance with a person’s claim) are the main tasks involving the use of biometrics. Until recently, identification and verification were accomplished on the basis of checking whether a person possesses a specific item (e.g. an ID card or a hardware key) or shares a specific knowledge (e.g. a password or a PIN). Both of these approaches have significant disadvantages, as required items and knowledge are subjected to the risk of being stolen, guessed, lost or forgotten [5].

One of the feasible answers to these threats is basing the process of checking the identity on the analysis of distinctive features of a human body, which can be measured strictly and expressed numerically [6]. Such feature is described by a term ‘biometric’, which is derived from Greek words ‘bio’ (life) and ‘metric’ (measure, measurement). Because each biometric is permanently connected with the person’s body, it cannot be lost or stolen and is always available for analysis. It seems to be the most natural possibility of confirming a person’s identity. Furthermore, the identification can be conducted even without the will or knowledge of a person, which can be useful in applications connected with crime detection and intelligence. A biometric can be, however, imitated and designers of automatic recognition systems should take this into consideration.

The process of identification requires comparing the measured biometric characteristic with every model stored in a biometric database of automatic recognition system and choosing the one that is most similar. Verification is less demanding, as it requires comparing the measured characteristic with only one model chosen from the database. The choice of the model from the database is based on a claim made by the verified person. The claim can be expressed, for instance, by inputting ID number or login. On the basis of the comparison, the similarity is determined. If its value exceeds a fixed bias, the identity of the claimant is confirmed. In the opposite case, the claim is rejected as a false one [6].

In general, biometrics can be divided into two main groups: behavioral and physical biometrics. A physical biometric is connected to the anatomy of a person’s body. A
behavioral biometric is dependent on person’s conscious or unconscious behavior or activity. The first group encompasses, among other things, fingertip pattern, iris and retinal scan, facial recognition, hand geometry, while the other includes analysis of signature dynamics, keystroke dynamics and vocal behavior [5].

Although voice is considered to be a behavioral biometric, it depends strongly on the anatomy of the vocal tract. According to the level of the complexity which biometric features are derived from, dependency on the anatomy will differ. One of the reasons why low-level features (i.e. short-term spectrum, linear prediction coefficients, cepstrum) are frequently used in speaker recognition is the fact that these are to a greater extent anatomically determined than high-level features (i.e. syntactical and lexical characteristics). These are thus more independent of person’s emotional state and are difficult to imitate. Moreover, these can be easily extracted and enable creating less memory-consuming speaker models [5].

2 Support Vector Machine

Choosing the human speech as the basis of the verification or identification places demands on the classifier used during the decision process. As the human voice is highly changeable in the long-term, as it depends on the time of day and can be influenced by a sickness, a classifier should be both highly discriminative and able to generalization. Because anticipated applications of automatic speaker recognition systems include internet banking and phone commerce, the classifier should also be computationally efficient. Support Vector Machine, which was introduced in 1992 by Boser, Guyon and Vapnik [1], seems to be the promising solution of this problem.

Support Vector Machine is an algorithm for learning linear classifier in the kernel-induced vector spaces. The obtained solution of the classification task is computationally efficient and optimal in terms of insights from the field of generalization theory.

2.1 Kernel-induced feature spaces

A linear classifier is capable of separating feature vectors into two classes $\omega_1$ and $\omega_2$ only if the classes are linearly separable [3]. In other words, there has to exist a hyperplane described by normal vector $\mathbf{w}$ and translation $b$ for which the following condition holds true:

$$
\begin{align*}
\forall x_i, \omega_i \langle \mathbf{w}, x_i \rangle + b > 0, \\
\forall x_i, \omega_i \langle \mathbf{w}, x_i \rangle + b < 0,
\end{align*}
$$

where $x_i$ are elements of a learning set $S$ and $\langle \mathbf{w}, x_i \rangle$ symbolizes the dot-product of vectors $\mathbf{w}$ and $x_i$. The decision function and the classifying function can be defined as follows [3]:

$$
\begin{align*}
f : X \rightarrow \mathbb{R}, & \quad f(x) = \langle \mathbf{w}, x \rangle + b, \\
h : \mathbb{R} \rightarrow \{-1,1\}, & \quad h(f(x)) = \text{sgn}(f(x)).
\end{align*}
$$

If such partition of the feature space cannot be found, solving the given problem by the straightforward use of a linear classifier is impossible. Unfortunately, in most of real-world situations vectors from learning set cannot be separated by any linear hyperplane in the input space $X$. Furthermore, if linearly separable classes are situated near to each other, the influence of various noise or measuring inaccuracies can make the given problem impossible to be solved by a linear classifier.
If the original learning algorithm can be transformed into so-called ‘dual algorithm’, in which the decision function includes only the dot-products of pairs of input vectors, kernel methods can be used to obtain a linearly separable problem. The foundation of this approach is an attempt to transform the original input space \( X \) by applying a non-linear mapping \( \Phi \) (Fig. 1). The linear classifier can operate in the new space, which is called the feature space \( F \) and can have higher dimensionality when compared to the input space. Owing to this, there is always a possibility of finding such mapping that the given problem will become solvable to the linear classifier, even though for some cases it may be necessary to use a powerful kernel in order to achieve this [3].

![Fig. 1 Using a non-linear mapping \( \Phi \) to transform the input space \( X \) into the feature space \( F \), in which the given problem becomes linearly separable and solvable for a linear classifier.](image)

However, if the dimensionality of the feature space becomes comparable to the cardinal number of the learning set, each feature vector can become linearly independent from other vectors [4]. Despite being able of a correct classification of elements from the learning set \( S \), the obtained classifier will prove completely ineffective on the testing set \( T \), as it will not be able to recognize any general similarities between elements presented during a learning phase, regarding each vector as a distinctive sample.

It is obvious that operating directly in a high-dimensional space would entail a significant increase in both capacity of memory needed to store transformed input vectors and computational complexity of the learning process. However, it is not necessary to use vectors from the feature space, as kernel methods exploit the possibility of calculating dot-products in feature spaces by determining the value of specific function defined in the original input space [4]. In order to achieve this, a specific function which satisfies conditions included in Mercer’s theorem [3] has to be found.

The feature space connected to the given kernel function is called a ‘kernel-induced space’. Because it is possible to calculate dot-products of feature vectors in the original input space, the dimensionality of the feature space is of no concern and can be even infinite. The most frequently used kernel functions are described by following equations [5]:

\[
K(x, z) = \langle x, z \rangle^d, 
\]

\[
K(x, z) = \exp\left(-\frac{1}{2\sigma^2}\|x - z\|^2\right),
\]

\[
K(x, z) = \tanh(\beta_0 x^T z + \beta_1).
\]
The kernel-induced space which is set by the equation (4) is called a polynomial kernel function has dimension equal to \(d\), and spaces set by equations (5,6) have infinite dimensionality and are referred to as radial and perceptron kernel functions, respectively [5].

Albeit the increase of the feature space dimensionality does not result in the rise of computational complexity of pattern classifier, it has other negative effects [3]. Firstly, the higher the dimensionality of the kernel-induced space \(F\) is, the more sparsely it is filled with the feature vectors belonging to the given learning set \(S\). In case of learning algorithms which result is ambiguous or which have high capatiy (e.g. perceptron algorithm), it can lead to receiving an ineffective classifier. On the other hand, using a more numerous set results in increasing both the duration of the learning phase and the amount of memory needed to store learning data. Moreover, the number of vectors in the learning set should be increased exponentially in relation with the rise of the feature space dimensionality – this objection is frequently referred to as ‘the curse of dimensionality’ [3].

2.2 Structural risk minimization

Even though for each classification task a kernel can be found which would be powerful enough to make the given problem solvable by a linear classifier, it often is detrimental. The learning set may include few elements which are not representative of the entire class (Fig. 2), for instance due to some measurement inaccuracies [3]. If the classifier is required to find a decision function which will give correct result for each vector from a learning set, the obtained solution will not be optimal in terms of structural risk in most cases. In other words, although the classification of vectors from the learning set will be correct (and therefore the empirical risk will be minimized), the classification of vectors from the testing set will be influenced by the presence of unrepresentative vectors (and so the structural risk will not be minimized). Vectors which have to be neglected during the learning of the classifier are called ‘outliers’ [3].

Using the structural risk minimization criterion as the grounds for rating the performance of the classifier allows one to obtain a reliable classifier and helps to avoid both obtaining unnecessarily complex classifying function and using excessively powerful kernels [3].

2.3 Maximal margin and soft margin Support Vector Machine

The simplest model of Support Vector Machine is the maximal margin SVM, which can classify only the data that are linearly separable [1].

As there is a direct relation between a functional margin, a geometrical margin and the normal vector \(w\) of the separating hyperplane [3], the problem may be stated as the task of
minimizing the norm of the normal vector $\mathbf{w}$ under the constraint that for each feature vector $\mathbf{x}_i$, the value of functional margin has to be greater or equal to 1.

$$\minimize_{\mathbf{w}, b} \langle \mathbf{w}, \mathbf{w} \rangle$$

subject to $\forall x_i \in S$, $y_i (\langle \mathbf{w}, \mathbf{x}_i \rangle + b) \geq 1$ \hspace{1cm} (5)

It is equivalent to maximizing the geometrical margin of the obtained separating hyperplane [3], which is equal to

$$\gamma = \frac{1}{\|\mathbf{w}\|_2}.$$ \hspace{1cm} (6)

The solution of the above problem can be obtained by exploiting a so-called ‘quadratic programming’ technique [3]. At first a primal Lagrangian that incorporates both minimized expression and linear inequality constraints $g(\mathbf{w})$ or equality constraints $h(\mathbf{w})$ has to be found, for instance

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2} \langle \mathbf{w}, \mathbf{w} \rangle - \sum_{i=1}^{l} \alpha_i [y_i (\langle \mathbf{w}, \mathbf{x}_i \rangle + b) - 1],$$ \hspace{1cm} (7)

where $l$ is the number of linear inequality constraints and $\alpha_i$ corresponds to Lagrange multipliers.

Inequality and equality constraints determine the feasible region $R$ in a domain $\Omega \subset \mathbb{R}^n$ which the optimization convex problem is defined on. It can be denoted by

$$R = \{ \mathbf{w} \in \Omega : g(\mathbf{w}) \leq 0 \land h(\mathbf{w}) = 0 \}.$$ \hspace{1cm} (8)

If the optimal normal vector $\mathbf{w}^*$ lies in the interior of the feasible region, all inequality constrain are called ‘inactive’, as they do not satisfy the condition

$$g_i(\mathbf{w}) = 0.$$ \hspace{1cm} (9)

In the opposite case, every inequality satisfying the above condition is called ‘active’ (or ‘tight’). The vectors for which respective equalities are tight are called ‘support vectors’ (Fig. 3), as only these decide about the position of the separating hyperplane and appear in the obtained classifying function [3]. As their number is usually significantly lower than the cardinal number of the learning set, the solution is sparse and computationally efficient, which makes it useful in applications connected with identification.

Fig. 3 Examples of decision boundaries obtained in the result of using the perceptron algorithm (ambiguous solution), the maximal-margin SVM and the soft-margin SVM with slack variables

The soft margin SVM classifier is the model which can be used to solve classifications problem that are not linearly separable, thus it is useful in real-world situations and applications. It was introduced in 1995 by Cortes and Vapnik [2]. Inequality constraints of this classifier are slightly loosen for some points (Fig. 3) by the introduction of so-called
slack variables’, denoted by $\xi_i$. These variables are subject to minimization as they added to the constraint’s part of the primal Lagrangian. For example, the 2-norm soft margin SVM Lagrangian [3] can be expressed by

$$L(w, b, a) = \frac{1}{2} \langle w, w \rangle + \frac{C}{2} \sum_{i=1}^{l} \xi_i^2 - \sum_{i=1}^{l} a_i \left[ y_i \left( \langle w, x_i \rangle + b \right) - 1 + \xi_i \right], \tag{10}$$

Slack variables make possible that several points will lie in the margin region or even in the region corresponding with the other class.

By introducing other constraints, various SVM-based classifiers may be obtained [3,5]. For instance, the elimination of vectors for which values of respective Lagrange multipliers were comparably large can be an example of data cleaning, as the larger the multiplier value, the more difficult classification of the given feature vector is. Also the maximal value of Lagrange multipliers can be constrained by using the 1-norm soft margin classifier. Finally, the optimization can be subject to the number of support vectors, which enables the direct control of the complexity of the decision function.

3 Summary

Support Vector Machine is an algorithm for learning linear classifiers that enables the use of kernel methods. The obtained classifier is optimal in terms of generalization abilities and provides a sparse and computationally effective solution of the classification problem by exploiting optimization methods of Lagrangian theory, further developed by Kuhn and Tucker. As various optimization criteria can be introduced, the functioning of the SVM classifier can be adjusted to the examined problem within the wide range.

4 References

Application of PCA algorithm for DNA microarrays gene expressions

Łukasz Ćmielowski, Patryk Walawender
Silesian University of Technology
cmielowski@tlen.pl, walek24@tlen.pl

Abstract: This article examines the application of PCA algorithm for DNA microarrays gene expressions. DNA microarrays technology has made thousands of genes expressions data available for various analysis. Because of huge amount of data it was necessary to provide mathematical and computation tools to help efficiently handle such large datasets. The PCA is a statistical and computational method for simplifying complex data sets. The PCA method has already found application in fields such as face recognition and image compression, and is a popular method for matching patterns in high dimension data sets. The aim of the article is to show how PCA algorithm can be applied for DNA microarrays gene expressions by reducing the dimension of datasets and thanks to that allowing for easier data visualization, noises filtering and simplifying the subsequent analysis.

Keywords: gene expression, DNA microarrays, principal component analysis (PCA), pattern discovery.

1 Introduction

What DNA micro-arrays are really after are the patterns of expression across multiple genes and experiments. Multi-gene expression patterns could characterize diseases and lead to new precise diagnostic tools capable of discriminating, for instance, different kinds of cancers. Principal component analysis (PCA) is a widely used statistical data analysis technique that can be viewed as: 1) a method to discover or reduce the dimensionality of the data; 2) a method to identify new meaningful underlying variables; 3) a method to compress the data; and 4) a method to visualize the data. Array data is inherently high-dimensional; hence methods that try to reduce the dimensionality of the data and/or lend themselves to some form of visualization remain particularly useful [6].

1.1 Visualization of input data set

The input data set contains tissue samples of 4 patients with endometrial cancer and tissue samples of 3 patients in which neoplasia was not histopathologically diagnosed (control material) [1]. The dimensions of input data are as follows: 120 genes for each 7 samples (4 with endometrial cancer and 3 control ones). Because the main aim of the paper is present the PCA algorithm and its application to DNA microarrays gene expressions we will focus more on computer aspects of problem than on biological interpretations.
As we can imagine the analysis of plotted data (presented input data can be consider as simple/small data set comparing to other DNA microarrays data sets) is not a trivial task for computation not even mention “by hand” analysis. By computation problem we mean efficiency, time consumption, and making the data simple enough by dimension reduction for additional analysis.

2 PCA method

Principal Component Analysis is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. It is a linear projection method which defines a new dimensional space that captures the maximum information present in the original input data set by minimizing the error between the original data set and the reduced dimensional data set. Each principal component is defined such as to be orthonormal to all others and to maximize the information in the data that has not already been captured by the previous (lower) dimensions [2, 3, 7]. The PCA algorithm performs the following steps:

- input data
- mean subtraction from input data
- covariance matrix calculation
- eigenvectors and eigenvalues calculation of the covariance matrix
- choosing components and forming a feature vector (1). Eigenvectors are sorted according to their significance i.e. the one with the highest eigenvalue are first [2].

\[
\text{feature\_vector} = (\text{eig\_vector}_1, \text{eig\_vector}_2, ..., \text{eig\_vector}_n) \tag{1}
\]

Now is the moment when reduction of data dimension comes into. Each eigenvalue provides a measure of the proportion of the variance explained by the corresponding
eigenvector [6]. Basing on eigenvalues we can decide how many of eigenvectors of lesser significance can be ignored (left) during feature vector forming (1).

Fig. 2 Eigenvalues of input data set (7 samples)

Basing on eigenvalues we decide to ignore eigenvectors with indexes from 3 to 7, i.e. our feature vector will contain two eigenvectors with the highest eigenvalues.

- creating new reduced data set

\[ \text{final\_data} = (\text{feature\_vector}^T \ast \text{mean\_subtracted\_data}^T)^T \]  

After leaving five of the eigenvectors out, final data set (2) has reduced dimensionality to two, i.e. the new data is only in terms of the vectors that we decided to keep.

2.1 Information loss

The other main advantage of PCA is that once you have found patterns in the data, and you compress the data, i.e. by reducing the number of dimensions, without much loss of information [2].

Because of skipping some eigenvectors there always occurs some information loss. Figure 3 shows how information loss depends on number of eigenvectors chosen to calculation of PCs. The original input data can be retrieved according to following formula:

\[ \text{retrieved\_input\_data} = (\text{feature\_vector} \ast \text{final\_data}^T)^T + \text{subtracted\_mean} \]  

To measure the information loss depending on number of eigenvectors used in forming feature vector we provide the following simple measurement:

\[ \text{inf\_loss\_value} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} |\text{input\_data\_value}(i, j) - \text{retrieved\_data\_value}(i, j)|}{n + m}, \]  

where \( n \times m \) – final data size (2)
In other words, the $inf\_loss\_value$ is the mean of absolute difference between original data values and retrieved ones (4).

Fig. 3 Dependency between number of eigenvectors used in dimension reduction and information loss (for better display the minimum value of input data set and average one are marked by lines: ‘--’, ‘-.-’)

As we can see on figure 3 using all 7 eigenvectors allow us to retrieve original input data without information loss. Using only two eigenvectors in PCA scores calculation (see fig. 2) gave us $inf\_loss\_value = 0.2$, taking under consideration fact that we reduced our input data set from \(120 \times 7 = 840\) to \(120 \times 2 = 240\) it is acceptable loss.

2.2 Application of PCA to input data set

Fig. 4 Bars show the percentage of the total variance explained by each principal component, while the upper line shows the cumulative variance explained by the components
The principal components are sorted in decreasing order of variance, so the most significant components are always listed first. The first and second principal components explain the total variance of our input data set in ~ 97%. In this data, most of the important biological behavior is being captured in these two components, sending a message to take a closer look at them and their meaning in the context of the endometrial cancer (fig. 5)[5].

The first dominant eigenvectors can be associated with the discovery of important features or patterns in the data. In DNA micro-array data where the points correspond to genes and the axes to different experiments, such as different points in time, the dominant eigenvectors can represent expression patterns. For example, if the first eigenvector has a large component along the first experimental axis and a small component along the second and third axis, it can be associated with the experimental expression pattern "high-low-low" [6].

The paper [1] summarizes that changes in transcriptional activity of genes encoding proteins of mitogenic pathways induced by kinin receptors, such as 3CREB, 3KRAS, 1GNB1, 3cJUN, 1AC9, FOS, 2SRF, 1PRKACB, PLCB4, PRKARB2, may be taken into account in the discrimination between normal endometrium and endometrial cancer. Paper [1] also emphasizes that FOS and 3CJUN are the most interesting ones. Basing on that knowledge we can see that using PCA (fig. 5) the following genes: 3cJUN, 1AC9, FOS, PLCB4 have been “highlighted” as the ones worth further detailed inspection. Taking into account that no biological experiences, databases and previous works were used to eliminate not significant genes from the group of 120 genes the results of PCA although not perfect allows for gathering useful information for further analysis.

Fig. 5 Scatter plot of First vs. Second Principal component (the lower plot contains colored clusters)

The paper [1] summarizes that changes in transcriptional activity of genes encoding proteins of mitogenic pathways induced by kinin receptors, such as 3CREB, 3KRAS, 1GNB1, 3cJUN, 1AC9, FOS, 2SRF, 1PRKACB, PLCB4, PRKARB2, may be taken into account in the discrimination between normal endometrium and endometrial cancer. Paper [1] also emphasizes that FOS and 3CJUN are the most interesting ones. Basing on that knowledge we can see that using PCA (fig. 5) the following genes: 3cJUN, 1AC9, FOS, PLCB4 have been “highlighted” as the ones worth further detailed inspection. Taking into account that no biological experiences, databases and previous works were used to eliminate not significant genes from the group of 120 genes the results of PCA although not perfect allows for gathering useful information for further analysis.

The paper [1] summarizes that changes in transcriptional activity of genes encoding proteins of mitogenic pathways induced by kinin receptors, such as 3CREB, 3KRAS, 1GNB1, 3cJUN, 1AC9, FOS, 2SRF, 1PRKACB, PLCB4, PRKARB2, may be taken into account in the discrimination between normal endometrium and endometrial cancer. Paper [1] also emphasizes that FOS and 3CJUN are the most interesting ones. Basing on that knowledge we can see that using PCA (fig. 5) the following genes: 3cJUN, 1AC9, FOS, PLCB4 have been “highlighted” as the ones worth further detailed inspection. Taking into account that no biological experiences, databases and previous works were used to eliminate not significant genes from the group of 120 genes the results of PCA although not perfect allows for gathering useful information for further analysis.
3 Conclusions

Although PCA’s primary focus is not data clustering, projection on to lower dimensional spaces associated with the top principal components can be helpful in revealing and visualizing the presence of clusters in the DNA microarrays gene expressions data (fig. 5). These projections however must be considered carefully since clusters present in the data can become hidden during the projection operation [6]. Although PCA is a very useful method, it is only one way of analyzing DNA microarrays data which should be complemented by other techniques, and in particular by methods whose primary focus is data clustering. In our future work we would like to make an attempt of combining the PCA algorithm with techniques based on fuzzy logic; we see here a lot of possibilities for improving efficiency and accuracy of fuzzy logic method by combining it with data reduction focused techniques.

4 References

The optimization of measuring method of blood glucose from dry bar code tests.

Rafal Doniec
Silesian University of Technology
e-mail: rafal.doniec@telekomunikacja.pl

Abstract: The optimization of measuring method of blood glucose from dry bar code tests. The comparison accessible on Polish market of dry tests to measurements blood glucose. And base deployment and comparison of measuring method based on regression on traditional methods of curve fit of concentrating glucose in blood.

1 Introduction and profile of reflection in the bar code tests

Spectral reflection technology reduced the cost of investigation by simplifying analytic procedure, shortens the time of examination and makes patient examining on spot. Size reduction of measuring devices and elimination of the chemical reagents in traditional methods are the additional advantages of the method.

Disadvantages of the method are deferens’s between clinical results of measurements and results from spectral reflection technique. The spectral reflectance technique is subjective measurement, depending on visual color option of the suitably prepared belt and comparing it with scale. This way we received the half quantitative result it was necessary to enlarge an accuracy and repeatability of dry test. Therefore the first reflectation interpolators were worked out where the measurement of monochromatic radiation was used.

The main part of reflectation interpolator are scanner, reflectation converter, measuring or processing on electric signal of monochromatic radiation reflecting from altering color surface influence of measuring sample.

We can measure different biological sizes with this method but the measurement of blood sugar interests us the most. The first scanner has Reflectance Meter, produced by Ames Company which worked out the next three models: Eyetone, Dextrometer and Glucometer. The two last were based microprocessor engineering and had a measuring range from 0-400mg/dl as well as digital results display. Together with the growth of meaning and the dropping cost of microprocessors production the new micro technologies the eightieth it is: SSI, MSI, and LSI until ninetieth the ULSI or the Pentium technology it begun to work out enlarging of objectivity and accuracy measurement, scanners of dry tests.

1.1 Dry test accessible on Polish market

At present there accessible a range of scanners models differentancy from the technical, the measuring method, variety of measurments, or price on Polish market.
The most popular scanners are as follows: Accu-Chek, One Touch II, Profile, Glucometr Elite and Sure Step. The interest of the measurement of blood glucose rigor with using the reflectation scanners of dry test comes from 1971 when it was produced by Ames company the first device of this type called Reflectance Meter cooperating with Dextrostix electrochemical capillary belts came into being systematically new constructions of dry test such how: Glucostix, Glucofilm on a market. In Poland glucometrical scaners are refinanced therefore the increasing accessibility of these devised on market.

Most compiled models permit on remembering the last results, sending of data and straight matching simple statistical comparisons using PC equipment in such medias as Windows e.g. Accu-Chek Easy or devices including processors and used recently in PC equipment and suitably equipped permit on record and analysis of result from many months in popular installed PC programmers at present. However the most essential features characterizing scanners are: repeatability and stability of measurements. Every Glucometrr manufacturer adapts it to read a definite type of belts, using to read one or many parameters. We member to invasive methods carried out with the help of bar code tests analyses, are considered to be invasive methods connected with inconvenient operation puncturing. Its very essentional problem because with many diabetes the immunity of organism is weaker.

### 1.1.1 Other alternative methods of measurements of concentrating of contents of sugar in blood.

Elaboration appear on market presently substituting coulometer by electro chemical, as it make in devices Glucocard II from Arkray, Exac Tech Sensor from MediSense or GlucoWatch from Animas Technologies. There are meters of electric loads fabricated among electrodes on test strip thong for with glucose included in blood influences of reaction oxidize.

New methods emerge for collecting sample of blood completely also escaping operation invasion as for example, scanning with the aid of apparatus eye firm like in Spectrx from Boehringer Mannheim. Mentioned methods all highly, with exception last, they require employment specially strip preparation and also they are very high cost in daily application.

### 2 An optimization measurement of glucose concentration and comparison polynomial with wavelet regression method

With the aid of series of strip Accutrend and manner receive on orientating to that converter perform series measurement allowing act test dry Accutrend and formulating of optimal algorithm on this base procedure and dependences related with measurement of concentrating glucose in blood. With the aid of series of strip Accutrend and reflectometer we orientating to that how series measurements perform. And formulating of optimal algorithm on this base procedure. Dependences related with reflectation measurements of blood glucose behind assistance of dry test Accutrend Glucose

For determination of mathematical dependence, wanted for formulating concrete project foundation, it carries proper research:

- Exact analysis of structure of strip Accutrend, with its particular verification of strip work;
- Conducting manner of measurement with laboratory reflectometer, there be exit base for elaboration of measuring procedure;
- Determination of stability of measuring parameter in time
- Determination on base of research of dependence in function of wave length of monochrome light radiation for different value glucose concentration on Accutrend strip.
Determination of dependence reflectance for Accutrend strips function of concentrating glucose
Optimization of measurement research.
Place centrally measurement field is main part of strip Accutrend Glucose. Blotting paper is
impregnated component of measurement field specially, drawn reacts to it with reached
glucose included in blood mixture chemically.
For determination of mathematical dependence, wanted for formulating exemplars describing
dependence concentrate solution but reflectation proper research of chemical reactions in field
carry getting test dry take advantage Accutrend, laboratory reflektometer, recorder and digital
(numeric) voltmeter. In about give length of wave fabricate radiation reflektometer, entering
to sphere and from surface of measuring strip reflection, it has been processed on electric
signal. Current electric signal, from exit of photodiode, current is subjected transformation on
signal in converter tense to voltage. We gauged this signal with the aid of digital voltmeter.
We have measured be for flow of radiation reflected signal has been measured proportional
which voltage and it can be converted on according to exemplar (1)

\[
R = \frac{U_{P(on)} - U_{P(off)}}{U_{W(on)} - U_{W(off)}} \tag{1}
\]

\(U_{P(on)}\) - level of signal at measurement of researched sample gotten, at including source of
radiation;
\(U_{P(off)}\) - level of signal at measurement of researched sample gotten, without source of
radiation;
\(U_{W(on)}\) - level of signal at measurement of researched pattern gotten, at including source of
radiation;
\(U_{W(off)}\) - level of signal at measurement of researched pattern gotten, without source of
radiation;

At definition of dependence in function waves of length of monochrome radiation fundamental problem dry strip Accutrend Glucose, for different in blood glucose, it was about
different samples of blood received glucose concentrating. This problem with sample blood
solves by substituting blood glucose that pattern glucose solution. Measuring procedure
crossed to manner following prepared introduce strip for lecture for measuring cell.
Where at measurement of researched sample \(U_{P(on)}\) - signal gotten , at including source of
radiation; at measurement of researched sample - signal gotten , at excluded in different ratio
of wave.
The next measurement step \(U_p\) at definition of dependence in function of wave’s length of
monochrome radiation fundamental strip Accutrend Glucose, and different concentration
glucose.
Time of measurement has been defined on 20s.
Received results have allowed drawing characteristics of reflectation, in function of waves
length. Correlation among results of measuring series amount to computational 0,9.
Wavelet analysis tends to find a prototype of function called mother wavelet. The
-glucose concentration measurements process can be illustrated as a combination of
elementary vectors. Therefore it is possible to express each function by the combination of
elementary vectors consisting of baseline elements \((1,0)\) ,\((0,1)\) and other vector’s element as
a combination \((1,0)\) and \((x,0)\) with \((0,1)\) and \((0,y)\) of linear functions. There is an additional
characteristic feature of baseline vector i.e. it is its orthogonallity.
Both scaling and translating of mother wavelet allows defining the orthogonal base. Then
from the orthogonal base a suitable matrix of mother wavelets can be created. Its main
advantage is the possibility of their similarity recognition. Such similarity recognition requires n-dimension translations in the n operations. Recent investigations referring to the development of glucose-insulin dependence modeling use several feature of the real pancreas. Presented solution, based upon application of wavelet transform has created a model imitating glucose concentration dependence. The linear dependence between reflectance and glucose concentration has been confirmed based on measurements and modeling.

3 An optimization measurement of glucose concentration and comparison polynomial with wavelet regression method

The system based on wavelet transform has potential to facilitate real-time concentration glucose monitoring and is based for construction of device for blood glucose throughout the clinical range. Recent researches on development of modeling glucose-insulin dependence are creation of close regulating-measuring system, and functional imitating work pancreas. To present solution which based on wavelet transform are model of measurement imitating dependence reflectation from glucose concentration. Receive dependence due to results of measurements concentrated simulate dependence reflectation from glucose concentration it receive linear dependence. The system based on wavelet transform has potential to facilitate real-time concentration glucose monitoring and is base for construction of device for blood glucose throughout the clinical range.

References

Citations referred in text by numbers in square brackets. All references listed at the end of the paper after the heading “References”. Numbers not indented, text shifted 1,5 cm from the left border of the page.
Statistical and Hybrid Modeling for Digits Recognition in Romanian Language

Corneliu Octavian DUMITRU\textsuperscript{1}, Inge GAVAT\textsuperscript{1}

\textsuperscript{1}Politehnica University Bucharest, Romania
e-mail: odumitru@alpha.imag.pub.ro

Abstract: This paper presents our work concerning speech recognition by statistical and hybrid modeling. The statistical modeling is realized by using Hidden Markov Models (HMM); the hybrid modeling (HMM-MLP) is realized by using Hidden Markov Models and Multi-Layer Perceptron (MLP). Trying to reduce the limitation effects of HMMs (low discrimination between models) and because the MLP have problems with accommodating time sequences like speech, a combination of a HMMs with a MLP could be a good idea. We used for speech parameterization 12 mel-frequency cepstral coefficients (MFCC). In a simple task of digit recognition, the recognition rate in the case of HMMs is lower than in the case of HMM-MLP with approximate 2.6%. For the HMMs, recognition rate can be improved extending the number of parameters (add their first and second order variation), extending the number of pronunciations and the number of speakers and optimizing the speech processing methods.

1 Introduction

Analyzing the speech recognition world, the leading position in realized recognizers is detained by the HMMs. Some of the many advantages of HMMs, which ensure their leading position in this moment, are:

- a very solid mathematical support of the training and evaluation algorithms;
- easy adaptation to the signal temporal dynamic without special warping methods;
- application in speech recognition systems based on the pattern recognition principle but also in expert systems.

The main disadvantage of HMMs consists in the low discrimination between models, caused by the training, based on a maximum likelihood (ML) strategy for the re-estimation of the model parameters.

Trying to reduce these limitation effects of HMMs, we chose an alternative; combining the two approaches (HMMs and MLPs) and we obtained a hybrid system.

The remainder of the paper is structured as follows: in chapter 2 is dedicated to Hidden Markov Models (HMM). Chapter 3 is dedicated to hybrid system (HMM-MLP). Database and experimental results are exposed in chapter 4. Conclusions close the paper.

2 Hidden Markov Models

HMMs are finite automates, having a given number of states; passing from one state to another is made instantaneously at equally spaced time moments. At every pass from one state to another, the system generates observations, two processes are taking place: the transparent one, represented by the observations string (feature sequence), and the hidden one, which cannot be observed, represented by the state string.
2.1 Left-right model
In Fig. 1. is represented the left - right model (Bakis), which is considered the best choice for speech. For each symbol, such a model is constructed; a word string is obtained by connecting corresponding HMMs together in sequence [1].

Fig. 1. Left - right model

2.2 HMMs Problems
Applying HMMs, there are three main problems:
- The first problem is the evaluation one, in the recognition phase, applied to the trained model. Given this model and the features sequence, we have to analyze if the sequence is produced by the given model. The probability to produce an observation sequence with a Markov model is calculated by the “forward” and “backward” algorithm.
- The second problem is about establishing the correct state sequence. The “Viterby” algorithm is one of the most used algorithms for this purpose.
- The third problem is the dedicated to the training, the parameter optimization of the model to describe as good as possible the observation sequence. Training allows optimal adaptation of the model parameters to the training set of data by re-estimating them. The “Baum-Welch” algorithm is the most used parameter re-estimation algorithm.

2.3 Algorithm
There are four important stages in the whole recognition process:
- the data preparation;
- the training of the models;
- the testing of the models;
- the final evaluation.

The data preparation is the first stage in the recognizer development. In order to build the HMMs, the set of corresponding speech data files and their associated transcriptions are required.

Before the training stage, the database must be converted into the parametric form; our first choice was the mel-frequency cepstral coefficients.
Transcriptions will also need preparing [8]. Typically the labels used in the original source transcriptions will not be exactly as required, for example, because of differences in the phone sets used. For our experiments it was necessary the labels to be context-dependent.

The training of the models is the second step of the system building; therefore is necessary to define the topology required for each HMM by writing a prototype definition. The system allows HMMs to be built with any desired topology. HMM definitions can be stored externally as simple text files and hence it is possible to edit them with any convenient text editor.

The purpose of the prototype definition is only to specify the overall characteristics and topology of the HMM. The actual parameters will be computed later by training. Initial values for the transition probabilities must be given and fortunately, the training process is very insensitive to these. An acceptable and simple strategy for choosing these probabilities is to make all of the transitions out of any state equally likely.

The actual training process takes place in the following steps:

- Allocate initial values and reset the accumulators for all parameters of all HMMs.
- Get the next training data. Construct a composite HMM by joining in sequence the HMMs corresponding to the symbol transcription of the training data.
- Calculate the “forward” and “backward” probabilities for the composite HMM.
- Use the “forward” and “backward” probabilities to compute the probabilities of state occupation at each time frame and update the accumulators.
- Repeat the step until all training data have been processed.
- Use the accumulators to calculate the new parameter estimates for all of the HMMs.

These steps can then all be repeated as many times as necessary to achieve the required convergence.

The testing of models. The system provides a single recognition strategy by use of the “Viterbi” algorithm. The recognition stage takes as input the data, a dictionary defining how each data is pronounced. The corresponding set of HMMs is allocated. Recognition can then be performed on either a list of stored speech files or on direct audio input.

The final evaluation. Once the HMM-based recognizer has been built, it is necessary to evaluate its performance. Using it to transcribe some pre-recorded test data and match the recognizer output with the correct reference transcriptions usually does this. This comparison is performed by the system with uses of the dynamic programming to align the two transcriptions and then count substitution, deletion and insertion errors [9].

3 Hybrid System: HMM-MLP

3.1 Structure & algorithm

The HMM-based speech recognition methods make use of a probability estimator, in order to approximate emission probabilities \( p(x_n / q_k) \), where \( x_n \) represents the observed data feature, and \( q_k \) is the hypothesized HMM state. These probabilities are used by the basic HMM equations, and because the HMM is based on a strict formalism, when the HMM is modified, there is a great risk of losing the theoretical foundations or the efficiency of the training and
recognition algorithms. Fortunately, a proper use of the MLPs can lead to obtain probabilities that are related with the HMM emission probabilities [2], [3].

In particular, MLPs can be trained to produce the a posteriori probability $p(x_n/q_k)$, that is, the a posteriori probability of the HMM state given the acoustic data, when each MLP output is associated with a specific HMM state. Many authors have shown that the outputs of an ANN used as described above can be interpreted as estimates of a posteriori probabilities of output classes conditioned by the input, so we will not insist on this matter, but we will mention an important condition, useful for finding an acceptable connectionist probability estimator: the system must contain enough parameters to be trained to a good approximation of the mapping function between the input and the output classes.

Thus, the a posteriori probabilities that are estimated by MLPs can be converted in emission probabilities by applying Bayes' rule (1) to the MLP outputs:

$$\frac{p(x_n/q_k)}{p(x_n)} = \frac{p(q_k|x_n)}{p(q_k)}$$

That is, the emission probabilities are obtained by dividing the a posteriori estimations from the MLP outputs by estimations of the frequencies of each class, while the scaling factor $p(x_n)$ is considered a constant for all classes, and will not modify the classification.

This was the idea that leads to hybrid neuro-statistical methods, that is, hybrid MLP-HMM methods (Fig. 2), applied for solving the speech recognition problem [5], [6].

![Fig. 2. Schematic diagram of a hybrid system.](image)

4 Database and Experimental Results

We compare the performance obtained in digit recognition task with hidden Markov models (HMM) and with the hybrid system (HMM-MLP) for unrolled and enrolled speaker [4].

The database contains speech data from 9 speakers (6 male speakers (MS) and 3 female speakers (FS)) for 9 digits in Romanian language (unu, doi, trei, patru, cinci, şase, şapte, opt, nouă). We excluded the last 3 speaker (2 male speaker – speaker 9 (sp.9) and speaker 7...
(sp.7)) and one female speaker – speaker 8 (sp.8)) from the database and used them for the testing.

The audio files contain signals sampled with 16KHz, 16 bit, mono, recorded with a desktop microphone, in a laboratory environment.

The digit parameters were extracted by cepstral analysis, in form of 12 mel-frequency cepstral coefficients (MFCC).

- HMM: for each digit we constructed a hidden Markov model. The speech data was parameterized using cepstral analysis in form of 12 MFCCs.
- HMM- MLP: the system consists of 9 hybrid models corresponding to 9 digits. Each hybrid model is made of 5 states, each state being associated with one output node of the MLP [7]. The MLP has one hidden layer (100 nodes), and the input layer consisting of 12 nodes [3].

The test results, for HMM and MLP-HMM for different training (for example: Tr = 9 - number of speaker used for training = 9) and different system (HMM or HMM - MLP) are given in the Table 1 and Fig. 3.

<table>
<thead>
<tr>
<th>Type: statistical / hybrid</th>
<th>Tr = 9</th>
<th>Tr = 8 FS</th>
<th>Tr = 7</th>
<th>Tr = 8 MS</th>
<th>Tr = 7 FS</th>
<th>Tr = 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMM-MLP</td>
<td>MS (sp.9)</td>
<td>100%</td>
<td>98.5%</td>
<td>91.5%</td>
<td>98.3%</td>
<td>89.9%</td>
</tr>
<tr>
<td>HMM</td>
<td>FS (sp.8)</td>
<td>100%</td>
<td>96.3%</td>
<td>88.9%</td>
<td>96.3%</td>
<td>87.7%</td>
</tr>
</tbody>
</table>

Table 1  Digit recognition rate for unrolled and enrolled speaker.

For the case of enrolled speaker, the error rate was:
- the same (0%) for hybrid modeling or statistical modeling,
- 1.5% for hybrid modeling and 3.7% for statistical modeling,
- 8.5% for hybrid modeling and 11.1% for statistical modeling.
For the case of unrolled speaker, the error rate was:
- 1.7% for hybrid modeling and 3.7% for statistical modeling,
- 10.1% for hybrid modeling and 12.3% for statistical modeling,
- 17.3% for hybrid modeling and 19.8% for statistical modeling.

The results obtained for unrolled speaker slightly lower than the results obtained for enrolled.

5 Conclusions

Because MLPs have problems with accommodating time sequences like speech and trying to reduce the disadvantage of HMMs, low discrimination between models, a combination of HMMs with MLPs in a hybrid system could be a good idea.

This hybrid system has been successfully applied for digit recognition within MLP was used as an a posteriori probability estimator of the HMM states [10].

Comparing the results obtained with hybrid system (HMM-MLP) and HMM we can see the results for HMM are slightly lower than the results for HMM-MLP with approximate 2÷2.6%.

For the HMM, recognition rate can be improved extending the number of parameters (add their first and second order variation), extending the number of pronunciations and the number of speakers and optimizing the speech processing methods.

6 References

Algorithms Evaluation for Remote Sensing Images Segmentation

Daniela Faur*, Mihai Datcu**, Inge Gavat*
*University Politehnica Bucharest-UPB, **German Aerospace Center-DLR

dfaur@electro.masuri.pub.ro, mihaidatcu@dlr.de, igavat@alpha.imag.pub.ro

Abstract. The paper presents the implementation of some algorithms for remote sensing image segmentation. New applications of remote sensing data require knowledge of the complicated spatial and structural relationships between objects within an image. This knowledge is ‘hidden’ in the image’s structure and must be ‘mined’ to retrieve meaningful spectral or polarimetric signatures or objects of higher-level abstraction, such as cities, roads, rivers, forests, etc. Algorithms for image segmentation like region growing, watershed transformation and mean shift procedure are presented and the results are compared.

1 Introduction

Interest in remote sensing image processing is continuously growing because the huge information content that can be extracted and used for different purposes like: flooding, mapping and forecast, land cover and land use change, environmental monitoring, cartography, meteorology etc. Mid-level processing on images involves task like segmentation – description of image into regions or objects - description of those objects to reduce them to a form suitable for computer processing and recognition of individual objects.[1].

An image segmentation is typically defined as an exhaustive partitioning of an input image into regions each of those considered to be homogeneous with respect to some image interest property e.g. intensity, color, texture [Jain at al 1995].

If the original image is described by the equation (1):

\[ P = \{ x_{ij}, i = 1,...,n_r, j = 1,...,n_c \} \] (1)

where \( x_{ij} \) represent each pixel of the original image with \( n_r \) rows and \( n_c \) columns, then the segmentation can be expressed as: \( S = \{ S_1,...,S_k \} \) with the \( l \) – th segment:

\[ S_l = \{ (i_{n_1},j_{n_2}),......,(i_{n_m},j_{n_m}) \} \] (2)

consisting of a connected subset of the pixel coordinates. Now two segments share any pixel locations ( \( S_i \cap S_j = \Phi, \forall i \neq j \) ) and the union of all segments covers the entire image.

Image segmentation algorithms are based of one of two properties of intensity values: discontinuity or similarity. In the first category the approach is to partition an image based on abrupt changes in intensity, such as edged in an image. The main approaches in the second category are based on partitioning an image into regions that are similar according to a set of predefined criteria. Thresholding, region growing, watershed transform and mean-shift are examples of methods in this category.
2 Region growing

The goal of region merging and region growing is to divide the domain $R$ of the image $I$ into $R_i$, $(i = 1..M)$ regions so that $R = \bigcup_{i=1}^{M} R_i$, $R_i \cap R_j = \emptyset$ if $i \neq j$ and $I$ satisfies a homogeneity criterion on each $R_i$.

Region growing approach has the following steps:
1) Start by choosing an arbitrary seed pixel and compare it with the neighborhood pixels.
2) Region is grown from the seed pixel by adding in neighboring pixels that are similar, increasing the size of the region.
3) When the growth of one region stops we simply choose another seed pixel which does not yet belong to any region and start again.

This whole process is continued until all pixels belong to some region. However starting with a particular seed pixel and letting this region grow completely before trying other seed biases the segmentation in favor of the regions which are segmented first. This can have several undesirable effects:
- Current region dominates the growth process (ambiguities around edges of adjacent regions may not be solved correctly).
- Different choices of seed may give different segmentation results.
- Problems can occur if the arbitrarily chosen seed lies on an edge.

Although region growing algorithms are very intuitive then can rarely be proven to converge to the minimum of the global cost function and the resulting regions may have noisy regions.

3 Watershed segmentation

In geography a watershed is a ridge that divides areas drained by different rivers systems. A catchment basin is the geographical area draining into river. The watershed transform applies these ideas to gray scale image processing in a way that can be used to solve a variety of image segmentation problems. Image data may be interpreted as a topographic surface where the gradient image grey levels represent altitude. Region edges correspond to high watersheds and low gradient regions interiors correspond to catchment basins. Catchment basins of the topographic surface are homogeneous in the sense that all pixels belonging to the same catchment basin are connected with the basin’s region of minimum altitude (grey level) by a simple path of pixels that have monotonically decreasing altitude (grey level) along the path. Such catchment basins represent the regions of the segmented image.

Fig. 1 One dimensional example of watershed segmentation: (a) Grey level profile of image data. (b) Watershed segmentation – local minima of grey level (altitude) yields catchment basins, local maxima define the watershed lines.
Two basic approaches are considered for watershed image segmentation:

**I)** The first one starts with finding a downstream path from each pixel of the image to a local minimum of image surface altitude. A catchment basin is then defined as the set of pixels for which their respective downstream paths end up in the same altitude minimum. While the downstream paths are easy to determine for continuous altitude surfaces by calculating the local gradients, no rules exist to define the downstream path uniquely for digital surfaces.

**II)** The second approach is essentially dual to the first one: instead of identifying downstream paths the catchment basins fill from the bottom. A hole is imagined in each local minimum and the topographic surface is emerged in water – water starts filling all catchment basins minima of which is under water level. If two catchemnt basins would merge as a result of further immersion a dam is built all the way to the highest surface altitude and the dam represents the watershed line.

### 4 Mean Shift algorithm

As it was discussed in [4] one approach for image segmentation is the classification of the features space. Arbitrarily structured features space can be analyzed only by non parametric methods since these method do not have embedded assumptions. Numerous non parametric clustering methods were described in the literature and they can be classified into two large classes: hierarchical clustering and density estimation. Hierarchical clustering techniques either aggregate or divide data based on some proximity measure. The hierarchical methods tend to be computationally expensive and the definition of a meaningful stop criterion for the fusion or the division of the data is not straightforward. The rational behind the density estimation based clustering approach is that the features space can be regarded as the empirical probability density function (p.d.f.) of the represented parameter. Dense region in the feature space thus correspond to local maxima of the p.d.f., the modes of the unknown density. Once the location of a mode is determined the cluster associated with it is delineated based on local structure of the features space.

An interesting approach to mode detection and clustering is based on the mean-shift procedure, proposed in 1975 by Fukunaga and Hostetler [21-comaniciu] but the advantages of employing a mean-shift type procedure in density estimation were only recently discovered. Kernel density estimation, known as Parzen window in pattern recognition literature, is the most popular density estimation method. Given n data points \( x_i, i = 1, 2, ..., n \) in the \( d \)-dimensional space \( \mathbb{R}^d \), the density estimator is given by:

\[
\hat{f}_{h,k}(x) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^{n} k\left( \left\| \frac{x-x_i}{h} \right\| \right).
\]  

(5)

where \( h \) is the bandwidth parameter, and \( c_{k,d} \) is a constant.

The first step in the analysis of a features space with the above mentioned density function \( f(x) \) is to find modes of this density. The modes are located among the zeros of the gradient \( \nabla f(x) = 0 \) and the mean-shift procedure is an elegant way to locate these zeros without estimating density. The gradient of the density estimator is given by:

\[
\hat{\nabla f}_{h,k}(x) = \nabla \hat{f}_{h,k}(x) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^{n} (x-x_i) k\left( \left\| \frac{x-x_i}{h} \right\| \right).
\]  

(6)
Let us consider the function:
\[ g(x) = k' \left( \frac{x}{h} \right) \]  
assuming that the derivative of the kernel profile \( k \) exists for all \( x \in [0, \infty) \), except for a finite set of points. Now using \( g(x) \) for profile \( G(x) \) is defined as:
\[ G(x) = c_{g,d} \left( \| x \|^2 \right) \]
where \( c_{g,d} \) is the corresponding normalization constant. Introducing \( g(x) \) into (6) yields:
\[ \hat{f}_{k,K}(x) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^{n} (x_i - x)g \left( \left\| \frac{x - x_i}{h} \right\|^2 \right) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^{n} \left[ \sum_{i=1}^{n} g \left( \left\| \frac{x - x_i}{h} \right\|^2 \right) \right] \left( \sum_{i=1}^{n} \frac{x_i}{h} \right) - x \]
where \( \sum_{i=1}^{n} g \left( \left\| \frac{x - x_i}{h} \right\|^2 \right) \) is assumed to be a positive number.

From (5) the first term is proportional to the density estimate at \( x \) computed with the kernel \( G \):
\[ \hat{f}_{h,G}(x) = \frac{c_{g,d}}{nh^{d}} \sum_{i=1}^{n} g \left( \left\| \frac{x - x_i}{h} \right\|^2 \right) \]
The second term is the mean shift:
\[ m_{h,G}(x) = \frac{\sum_{i=1}^{n} x_i g \left( \left\| \frac{x - x_i}{h} \right\|^2 \right)}{\sum_{i=1}^{n} g \left( \left\| \frac{x - x_i}{h} \right\|^2 \right)} - x. \]
i.e., the difference between the weighted mean, using the kernel \( G \) for weights and, the center of the kernel (window). From (9), (10) and (11) becomes:
\[ \hat{f}_{k,K}(x) = \hat{f}_{h,G}(x) \frac{2c_{k,d}}{h^2 c_{g,d}} m_{h,G}(x), \]
yielding
\[ m_{h,G} = \frac{1}{2} h^2 \frac{\hat{f}_{k,K}(x)}{\hat{f}_{h,G}(x)}. \]
The expression (13) shows that at location \( x \), the mean-shift vector computed with the kernel \( G \) is proportional with the normalized density gradient obtained with the kernel \( K \). Thus, the mean shift vector always points toward the direction of the maximum increase in the density.

Since the features space analysis can be applied to moderately higher dimensional spaces a joint domain representation is employed in the mean shift algorithm for image segmentation. The joint domain is the concatenation of the image spatial domain (pixel’s coordinates) and the range domain containing grey level, color or spectral information. For both domains the Euclidean metric is assumed. When the location and range vector are concatenated in the joint spatial-range domain of dimensions \( d = p + 2 \) their different nature has to be compensated by proper normalization. Thus the multivariate kernel is defined as a product of two radial
symmetric kernels and the Euclidean metric allows a single bandwidth parameter for each domain:

$$K_{h_x, h_y} = \frac{C}{h_x^2 h_y^2} k\left(\frac{\|x\|^2}{h_x^2}\right) k\left(\frac{\|x\|^2}{h_y^2}\right).$$

(14)

Where $x$ is the spatial part and $x'$ is the range part of a feature vector, $k(x)$ the common profile used in the both domain, $h_x$ and $h_y$ the employed bandwidths, and $C$ the corresponding normalization constant. In practice an Epanechnicov or a normal kernel always provides satisfactory results so the user has to set only the bandwidth parameter $h = (h_x, h_y)$ which determines the resolution of mode detection by acting on the kernel size.

5 Results and discussions

This section of the paper presents some results of the algorithms presented above applied on an optical remote sensing image. Also the analysis of land use satellite imagery depend heavily on the use of color, applying the region growing approach on this image yielded to no so bad results (fig.2.b).

Fig. a) An original IKONOS remote sensing sub-image; b) segmented image using region growing algorithm; c) results of watershed transform; d) segmented image with $h_x = 7, h_y = 6.5$, minimum region =10.
When the images are monochrome, region analysis must be carried out with a set of descriptors based on intensity levels and spatial properties. Another problem in region growing is the formulation of a stopping rule. Basically growing a region should stop when there are no pixels to satisfy the criteria for inclusion in that region. Additional criteria, besides intensity values, texture and color, to increase the power of a region growing algorithm utilize the concept of size, likeness between a candidate pixel and the pixel grown so far and the shape of the region being grown. The use of these types of descriptors is based on the assumption that a model of expected results is at least partially available.

Figure 2.c) depicts the results of a watershed segmentation using distance transform. The distance transform of a binary image is the distance for every pixel to the nearest non zero valued pixel. Note that some objects like the rivers were split improperly. This is called oversegmentation and is a common problem with watershed based method segmentation.

Finally the results of the mean shift algorithm are the most expressive for the imaged scene Fig. 2 d). We have used the Mean Shift based Image Segmenter. This soft package implements the image segmentation algorithm presented in [2] [7]. As described before this algorithm takes a feature (range) bandwidth, spatial bandwidth, and a minimum region area (in pixels) as input. The feature bandwidth may take on any real positive values; however, since the spatial subspace of the image data lies on the image lattice, the spatial bandwidth is constrained to all positive integer values. More specifically, the spatial bandwidth (r) defines a \((2r+1)(2r+1)\) spatial search window when computing mean shift. Minimum region is also constrained to all positive integer values.

Acknowledgement

Work in the frame of FP6-EC project -“Technology to Support Sustainable Humanitarian Crisis Management STREAM”.
The authors would like to thank to Robust Image Understanding Laboratory at Rutgers University for the Edison code.

References

Information Theory Measures for Changes Detection in Earth Observation Images

Daniela Faur*, Mihai Datcu**, Inge Gavat*
*University Politehnica Bucharest-UPB, **German Aerospace Center-DLR

dfaur@electro.masuri.pub.ro, mihaidatcu@dlr.de, igavat@alpha.imag.pub.ro

Abstract: This paper aims to present some information theory measures like entropy and Kullback-Leibler divergence to describe the information content in remote sensing imagery. Anomaly detection in Earth Observation images refers to the detection of irregularity in the scene that appears unlikely according to a probabilistic or physical model of the scene. The results of these measures denote how much an image scene corresponds to image information content. Optical remote sensing images, represented in feature space, are clustered using an unsupervised classification algorithm. Then, we infer the conditional model of the cluster structure in the data and we quantify image content using entropy measure.

1 Introduction

The global archives of remote sensing data are constantly increasing. First, the amount of operational remote sensing satellites is continuously growing, then, the spatial and spectral resolution of the new sensor are constantly improving due to technological advances, resulting in exponential larger data volumes. Finally, the data transmissions capabilities of modern satellites results in vastly increased data transfer.

Actually the end users utilize only a small percentage of the remote sensing data archived annually. Part of the reason is that most remote sensing applications depend on image processing carried out by experts. Even though the basic image processing such as geometric rectification and radiometric calibration is largely performed automatically for most remote sensing products, the conversion from image data to the thematic information required by end users is a constraining factor for the further proliferation of remote sensing applications. There is a need for developing automatic feature extraction methods that will make remote sensing information readily available to a wide range of end users. The “user” is considered as a broad category from professionals to the general public, from the environmental scientist to the analyst working in a county or a citizen seeking information through the Internet. For this reason EO data products must be exploitable in a user-friendly way through a clear understanding and description of the information and through digital data formats that are easy to handle.

New feature extraction methods are based on extracting and storing basic characteristics of image pixels and areas which are then selected by users as representative of the feature being searched for [1].
First, we use Shannon entropy as a well known measure for image information content. Another approach could be to find intermediate patterns as components for composite patterns and in later stages to build more semantic pattern. This can be done by ranking images by their information content in image regions or clusters in the image feature space, thus forming patterns with a certain semantic.

In our experiments we use about twenty remote sensing images acquired before and after the tsunami phenomenon from Indonesia. From space reasons this paper presents the obtained results only for two images. If in the first ones, before the event, it is easy to identify a certain structure of the imaged area like streets, buildings or forests, in the second ones, because the ocean washed out the surface of the earth, it is almost impossible to find any organized areas.

Anomaly detection using Earth Observation images refers to the detection of irregularity in the scene that appears unlikely according to a probabilistic or physical model of the scene, for example ammunition elements in a scene which is dominated by vegetation and soil or devastated areas in an urban zone.

2. Shannon entropy – based image content description

New trend is to organize large data sets using classification methods based on information content, arising hence the requirement to use image compression techniques. Information theory provides the mathematical framework to answer to the following question: is there a minimum amount of data that is sufficient to describe completely an image without loss of information?

Shannon’s “communication theory” deals only with objective information, treats the messages only from the point of view of engineering; it is limited to syntax, semantic is not considered [2]. Because the fundamental premise of information theory is that the generation of information can be modeled as a probabilistic process, we first develop a statistical model of the image generation process. As a simple example we can assume that the image was produced by an imaginary “8-bit grey level source” that sequentially emitted statistically independent pixels [3]. In this case, the source symbols are gray levels and the source alphabet is composed of 256 possible symbols. A method of estimating the information content is to construct a source model based on relative frequency of occurrence of the grey levels in an image under consideration. We suppose that the observed image can be interpreted as a sample of the behavior of the grey level source that generate it. Because the image is the only indicator of the source behavior, modeling the probabilities of the source symbols using the grey level histogram is reasonable.

An estimate of the entropy of the source can be computed using:

$$ H = - \sum_{i=1}^{N} P(a_i) \log P(a_i) $$

where $H$ is the average information per source output, $N$ is the number of source symbols and $P(a_i)$ is the probability of the event that source will produce symbol $a_i$.

Examining the obtained results for the two tsunami images is conclusive that entropy can be regarded as a measure to estimate image information content. Images that exhibit high entropy are very complex and it’s quite difficult to easily identify semantic patterns analyzing them. Low entropy images are much simple so that their information content is small. Notice
that the image’s entropy is decreasing after the phenomenon, suggesting small information content (tab I).

<table>
<thead>
<tr>
<th>Entropy</th>
<th>001</th>
<th>002</th>
<th>003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>13.69</td>
<td>14.56</td>
<td>15.63</td>
</tr>
<tr>
<td>After</td>
<td>9.15</td>
<td>13.11</td>
<td>13.88</td>
</tr>
</tbody>
</table>

Through the hierarchical image content characterization, the image retrieval system can be viewed as a composed communication channel. The image coding is effectuated by extracting its primitive features, i.e. color or spectral information, texture or geometrical parameters. Considering the information theory, which says that data processing cannot increase information, entails that each level in the hierarchical scheme is associated with a certain loss of information. The accuracy of communication, e.g. accessing a target image or a category of images as exploration results, depends on the assumed levels of image “coding”.

Fig. 1 Images of Banda Aceh city, Indonesia, before (first column) and after (second column) the tsunami phenomenon.

3. Image complexity valuation at cluster level

Majority of search applications are based on classification methods in features space. For optical remote sensing images features space is a 3-D space that contains the representation of pixels values R, G and B. Finding similarities between the patterns of the features space, image information content is reduced, compressed and the cluster is the one who provide information. Proceeding on this idea and on the fact that pattern recognition requires clusters recognition in different ways, this paper proposes a complexity valuation method at cluster level based on Kolmogorov entropy.

Kolmogorov generalized the notion of mutual information from information theory for the case of absolutely arbitrary continuous messages and signals. Starting with the mutual information he defines the $\varepsilon$-entropy of a random object. His problem formulation was: “Suppose source information yields a magnitude $X$ and we are required to code the information obtained with precision $\varepsilon$.” He defined the $\varepsilon$-entropy as the amount of mutual information of the magnitude $X$ and another signal $X'$ such that their joint probability density function is restricted to a certain family, dependent on the parameter $\varepsilon$ [4]. Kolmogorov gave an alternative interpretation to information theory from the algorithmic point of view.

Fig.2 It can be seen that after the water washed out the surface of the earth almost any structured form disappear creating anomalies regions.
In order to rank our image scenes, regarded as sub-images, by their information content, we begin by using a clustering algorithm like K-means on the image. The number of clusters is chosen according to rate distortion theory for image complexity determination [5] being the point at which the distortion curve levels off. Then we apply Kolmogorov entropy on clustered images for complexity valuation at cluster level. From the generated content-index the mutual information between image space I and class space ω (e.g. the K-means classified features space) can be computed as:

\[
I(I, \omega) = \sum_{i,k} p(\omega_i / I_k) p(I_k) \log \frac{p(\omega_i / I_k)}{p(\omega_i)}
\]

where \( p(\omega_i / I_k) \) indicates the posterior probabilities of signal classes given a certain sub-image \( I_k \) from the set (fig.2). Prior probabilities for signal classes and images are given by \( p(\omega_i) \) and \( p(I_k) \), respectively.

Table 2 The mutual information between image space and class space

<table>
<thead>
<tr>
<th>Mutual Information</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>0.264</td>
<td>0.336</td>
</tr>
<tr>
<td>After</td>
<td>0.144</td>
<td>0.286</td>
</tr>
</tbody>
</table>

The measures indicate the information transmitted from image data through feature extraction and unsupervised content-index generation (clustering) to the class space (tab. II). The obtained values depend on the estimated parameters (number of clusters) used for classification of every image using K-means algorithm. Using the algorithm presented in [5] for the first image number of clusters was 15 and for the second one 20, both before and after tsunami phenomenon.

The association between image space and class space can further be used to measure the complexity of the sub-images in the data set. Since the query performance of content-based image retrieval systems depends on the complexity of the data, analyzing the image database that is used for testing is rather important for evaluation. Similar to the method of [2, 7] that applies information theory to determine the complexity of image databases we measure the information between image space and class space based on Kullback-Leibler divergence. In comparison to mutual information, Kullback-Leibler divergence can be applied to determine the semblance of a scene with the image, being an objective measure. Kullback-Leibler divergence is given according to:

\[
D(p, q) = \sum_{i} p(\omega_i / I_k) \log \frac{p(\omega_i / I_k)}{q(\omega_i)}
\]

and can be interpreted as how much a single sub-image \( I_k \) is a typical mixture of the complete information content of the whole dataset or if the image is rather simple.

In order to determine anomalies regions we begin by identifying the interest regions of an image. Our approach is to consider every image as a date set by splitting it into equal size sub-images and to compute the semblance of a single sub-image in the entire set. The size of the sub-images is chosen according to image resolution and considering the size of the scene to be analyzed. Fig.3 a) shows the first image, divided in nine equal size sub-images.
Fig. 3  a) Optical remote sensing image A before the event. The image is split into nine images (250 x 250 pixels) from a1 to a9. b) KL divergence values for the image 001 before and after the tsunami phenomenon in Indonesia.

As it can be seen in fig. 3 b), both before and after the tsunami, the sub image a7 has the most different information content from the entire data set. Also the information content of the sub images a2 and a5 is similar with the information content of the whole set.

The most important observed changes are on the horizontal median of the image in a4, a5 and a6 while the images a1 and a9 have almost the same values. Fig. 4 a) shows the B image, revealing a coastal area. As the previous image this one too is divided in 9 sub-images, 256 x 256 pixels size.

Fig. 4  a) Image B before the tsunami phenomenon. b) KL divergence for image B before and after the tsunami.

In fig. 4 b) are depicted KL divergence values for the divided image B. As we expected to be, the different images from the data set are the ones containing mostly the forest (a2) or the urban area (a5) and the ocean in a9. By analyzing the results it is conclusive that the sub-image similarity with the set depends on the ability of the applied signal model to describe the
image content and to capture characteristic image structures. It is important to notice that Kullback-Leibler divergence was computed considering only the spectral information. Future work considers the introduction of texture model to complete the description of image information content.

3. Results and discussions

This paper presents an assessment of information theory methods for changes characterization in analysis of earth observation data. At the beginning, information measures settled in information theory namely Shannon’s entropy, mutual information and Kullback Leibler divergence are presented and then we applied them to evaluate information content in images before and after tsunami phenomenon. One relative approach for image content characterization is to assume a particular source model generating an image and compute entropy of the image based on that model. Mutual information between image space and class space can be regarded as an objective measure and indicates system’s ability to describe the image content. The composition of image semantics by combining clusters in the primitive image feature space is further used to consider the Kolmogorov approach. The algorithm has the following steps:

- Considering the remote sensing image as a data set we split it into equal size sub-images; the size of the blocks is chosen in respect to image resolution and user estimation about region of interest for anomaly detection.
- Entire image classification using K-means algorithm; the number of classes is chosen considering rate distortion theory for image complexity determination;
- Compute Kullback-Leibler divergence for a single sub-image similarity determination in the data set, in fact how much a scene is a mixture of the information content of the image. This approach might offer interesting applications in anomaly detection in image scenes, since it allows a comparison between the information contained in the scene and image information content.

Acknowledgement

Work in the frame of FP6-EC project -“Technology to Support Sustainable Humanitarian Crisis Management STREAM”.

References

Algorithms for inducing decision rules based on rough sets theory

Aleksandra Gruca
Silesian University of Technology
Aleksandra.Gruca@polsl.pl

Abstract: The rapid growth of the computer science technologies allows collecting huge databases with enormous number of information stored inside. To find and understand regularities in stored data sophisticated methods are needed. Interesting relations may be represented by decision rules which are intuitive form of representing knowledge. This paper presents several different algorithms for inducing decision rules based on the rough sets theory. The theory provides algorithms to analyze data tables where records are described by a vector of features. Particularly, if one or more attributes of this vector are decision attributes it is possible to generate a set of decision rules. Generated rules represent data patterns and can be used for description (knowledge discovery) or classification.

1 Introduction

The advent of information technology allowed easily collecting and storing data of previously unimaginable quantities. Nowadays, huge databases include possibly valuable and interesting information which are hidden beyond large data sets and cannot be extracted without using sophisticated methods.

Knowledge discovery is non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data [1]. One of the most well-known representations of such patterns are decision rules which represent relations found in the data in the natural form. Decision rules are considered to have a good ability of representing knowledge in the way easy to understand and interpret by human.

Rules are induced from training data set which includes information about objects described by a set of attributes. These rules are logical expressions used to classify new, unseen objects or to describe interesting relations found in the data. Generally speaking it is required that induced rules have both good classification and descriptive features.

There are many techniques used to induce decision rules. This article is focused on algorithms based on the rough sets theory introduced by Pawlak [2]. Rough sets theory is the powerful mathematical tool for data analysis which offers methods for efficient and universal decision rules discovery. Rules obtained by the rough sets theory methods usually have better abilities to recognize and classify objects than rules obtained by other methods. In real databases collected data is frequently imprecise and incomplete. Rough sets theory provides methods to deal with that vague and uncertain information what makes it very useful tool to analyze real data sets.

The first part of the article introduces very basic information on the rough sets theory while following parts present three algorithms which represent different approaches to the rule induction problem. Properties of generated rules very strongly depend on methods used for induction.

2 Basic notions

The information system $S$ is the couple $S=(U,A)$, where $U$ is a non-empty set of objects called universe and $A$ is a set of attributes. For each object $x \in U$ and attribute $a \in A$ we may define function $a:U \rightarrow V_a$, where set $V_a$ is the value set of the attribute $a$. 
In rough sets theory, analyzed data set is represented as an information system called decision table \( DT=(U,A\cup \{d\}) \) where \( U \) is a set of objects, \( A \) is a set of attributes called conditional attributes and \( d\in A \) is a distinguished attribute called decision attribute with domain \( D_d \). We may notice that decision attribute determines partition \( \{X_{d1},X_{d2}, \ldots ,X_{dk}\} \) of the set \( U \) with respect of value of the decision attribute \( d \).

For each \( d_i\in D_d \) the set \( X_i=\{x\in U: d(x)=d_i\} \) is called \( i \)-th decision class, and the partition \( U=\{X_1,X_2, \ldots ,X_k\} \) is called classification of the \( U \). Let \( S=(U,A) \) be an information system.

With every subset \( B\subseteq A \) there is an equivalence relation over universe \( U \) called \( B \)-indiscernibility relation, denoted \( IND(B) \), which is defined as follows:

\[
IND_S(B)=\{x,y\in U\times U: \forall a \in B \ (a(x)=a(y))\}
\] (1)

We say that objects from \( U \) satisfying the relation \( IND(B) \) are indiscernible by attributes of the \( B \). Having defined indiscernibility relation we may then define the notion of \( \text{reduct} \) which is the minimal subset \( B\subseteq A \) that \( IND_S(A)=IND_B(B) \). By \( RED_S(A) \) we denote the set of all reducts of the \( A \).

If we consider the decision table \( DT=(U,A\cup \{d\}) \), we may find it useful to define a \( \text{relative reduct} \) \( RED_S(A,d) \) which is the minimal subset \( B\subseteq A \) that \( IND_S(B)\subseteq IND_S(d) \).

As a decision rule we consider each expression of the form of:

\[
\text{IF} \ a_1\in V_{a1} \ \text{and} \ a_2\in V_{a2} \ \text{and} \ \ldots \ \text{and} \ a_n\in V_{an} \ \text{THEN} \ d=v
\] (2)

where \( v\in D_d \), \( \{a_1,a_2, \ldots ,a_n\}\subseteq A \) and \( V_{ai}\subseteq D_{ai} \), \( i=1,2,\ldots ,n \).

The left side of the rule is called conditional part while the right side is a decision part. Single expression \( a_i\in V_a \) is called a descriptor. The above expression is interpreted as follows: if attributes from the left side of the rule assume values from the set \( V_a \) then the decision attribute assumes the value \( v \).

In the literature a more general way of the rule notation \( \varphi \rightarrow \psi \) is common, where \( \varphi \) is the rule premise, and \( \psi \) is its consequent. This notation will be also used in the further part of the article.

Object \( x\in U \) matches the rule \( \varphi \rightarrow \psi \) if its attributes values satisfy the premise of a rule. The rule correctly classifies the object \( x \) if the object matches the rule and the decision given by the rule is the same as the value of decision attribute preset in a decision table. In this case we say that the object \( x\in U \) supports the rule \( \varphi \rightarrow \psi \).

The rule is true in the decision system if every object matching the rule is correctly classified by that rule. If the rule is true we may also say that the rule is accurate in the decision system, otherwise the rule is the approximate one.

Classification abilities of the obtained set of rules can be measured with the classification accuracy ratio. The decision table is divided into two parts: first one, called \( \text{training set} \), is used by the rule generation algorithm to learn the description of decision classes while the second one, called \( \text{test set} \), is used to verify whether rules are able to recognize objects correctly (to give objects the same decision value as the value preset in the decision table).

The classification accuracy is the ratio of the number of objects correctly recognized by rules to the number of all objects form the decision table.

Let \( DT=(U,A\cup \{d\}) \) be a decision table, \( DT_1=(U_1,A\cup \{d\}) \) be a training set, \( DT_2=(U_2,A\cup \{d\}) \) be a test set and \( U_1, U_2\subseteq U \). The classification accuracy of rules set generated on basis of the training table \( DT_1 \) and tested with the test table \( DT_2 \) is defined as follows:

\[
\text{accuracy}(D, DT_1) = \frac{\text{card}(\{u\in U_2 : f(u)=d(u)\})}{\text{card}(U_2)}
\] (3)

where \( f(u) \) is the value of decision given to the object by generated rules.
It is easy to notice that the classification accuracy assumes values form the interval [0,1]. The bigger value of the classification accuracy is, the better obtained rules might recognize unknown objects.

Usually so-called optimal decision rules are used for objects classification. In the rough sets theory the optimal decision rule in the decision system is the rule that is true in the system and removing any of its descriptors makes the rule no longer true in the system. More information about optimal decision rules can be found in [11][12].

Depending on the method used for rules generation we can divide rules generation algorithms into three different categories [10]:

- algorithms generating minimal set of rules,
- algorithms generating exhaustive set of rules,
- algorithms generating satisfactory set of rules.

Algorithms of the first group generate sets of rules as small as possible but still capable to describe input objects correctly. The opposite approach is represented by the second group – algorithms, which generate exhaustive sets of rules, are inducing all possible rules in the simplest form. Finally, the satisfactory set of rules is a set of rules generated according to some condition given by user.

Induced rules may be used for two different purposes: classification or knowledge discovery[4][10]. For classification tasks usually algorithms generating minimal set of rules are used while algorithms generating exhaustive and satisfactory set of rules are more popular in the area of knowledge discovery.

For the classification purpose we are interested in obtain rules both as general and as descriptive for data relations as possible and we usually do not care for size of the generated set. We want the obtained set of rules to be used for automatic classification of new, unseen objects which is a subject of machine learning studies.

For the data discovery purposes we expect that obtained rules reflect some previously unknown, interesting patterns in data. Sometimes we may want to present the generated set to an expert in order to better understand and explain dependencies between values of attributes and their associated decision classes. Each rule needs be evaluated separately as it may possibly represent an interesting pattern. The generated set should not be very extensive in order to simplify interpretation of the discovered knowledge.

## 3 Reduct based rules generation algorithm

Let $DT=(U,A\cup\{d^i\})$ be a decision table. It can be easily noticed that for each $R\in$RED$_S(A,d)$, the set of reduct attributes determines a decision value $d$. We may then simply use the relative reducts to generate a decision rule.

Algorithm generating the decision rules based on a decision reduct is presented below:

**input:** $S=(U,A\cup\{d^i\})$, RED$_S(A,d)$ – set of relative reducts

**output:** RUL($S$) – set of decision rules for $S$

**begin**

RUL($S$) = $\emptyset$

for each $R\in$RED$_S(A,d)$

for each $u\in U$

\[ r := \bigwedge_{a \in R} a_i = a_i(u) \rightarrow d = d(u) \quad /\!* create a decision rule r \quad /\!*/ \]

if $r \notin$ RUL($S$) then RUL($S$) = RUL($S$)$\cup\{r\}$

**end**

Rules generated by the classic method presented above have two essential disadvantages. They are too fitting to data and usually number of generated rules is very large. The first one...
means that rules do not recognize objects from outside of the decision table, therefore we can not use these rules to classify new, unknown objects. The second may be a problem if we want to use rules not for classification but for description of dependences which we are trying to find in the data.

Rules generated by the algorithm described above are not so-called optimal decision rules therefore cannot be used for objects classification in practice.

There is an extension of the algorithm which allows generating set of optimal decision rules, which can be found in [11][12]. However the set of rules generated by such an extended algorithm is also very large and therefore different approaches to rules generation problems are commonly used. One of such approaches is based on heuristic algorithms. Rules generated by such algorithms are able to recognize and correctly classify new objects while the set of generated rules is much smaller than the set of rules obtained by the classical method.

Heuristic algorithms are able to find the solution among all possible ones, but they do not guarantee that the solution is best. For this reason such algorithms are considered as approximate ones. The solution is usually close to the best and the result is found in fast and easy way. It may happen that solution is the best, but it can not be proved what makes the algorithm still called heuristic.

4 Sequential covering rules induction algorithm

Coverage of the decision system \( \mathbb{S} = (U, A \cup \{d\}) \) is a set of rules \( \text{RUL}(\mathbb{S}) \) of the following properties: for each object \( u \in U \) exists a rule \( r \in \text{RUL}(\mathbb{S}) \) classifying that object correctly. Coverage of the decision system \( \mathbb{S} \) is denoted as \( |\text{RUL}(\mathbb{S})| \).

One of the most well known approximate algorithms for the rule induction is the sequential covering algorithm. The idea of this algorithm is to learn one rule, eliminate all objects that apply to it and then repeat the whole process. Sequential covering algorithm may look as follows:

Input: \( (X_i \text{ – } i\text{-th decision class}, A \text{ – set of attributes}, U\text{-set of objects}) \)

Output: \( \text{RUL}(\mathbb{S}) \) - the coverage set of rules for the considered decision system.

Begin

\[ \text{RUL}(\mathbb{S}) = \emptyset \]

While (\( U \neq \emptyset \))

\[ r = \text{learn-one-rule}(X_i, A, U) \]

\[ \text{RUL}(\mathbb{S}) = \text{RUL}(\mathbb{S}) \cup r \]

\[ U = U \setminus [\text{RUL}(\mathbb{S})] \]

End

Return \( \text{RUL}(\mathbb{S}) \)

End

We assume that there is a procedure \textit{learn-one-rule}. The rule returned by that procedure should cover as many objects from the considered object set as possible. There are lots of methods of implementing the learn-one-rule procedure, some of them may be found in [12]. We may distinguish two main approaches to find the rule. First one assumes that at the beginning of the learn-one-rule procedure set of rule descriptors (left side of the rule) consist of all rule attributes. Then, descriptors are removed one by one and after each removal the rule is verified with desired quality. The second approach is based on the opposite idea. At the starting point left side of the rule is empty. Then descriptors are added. In both cases we assume that there is a decision assigned to the right side of the searched rule.

To decide when the process of adding or removing descriptors - depending on the learning rule method - should stop the rule quality measures are used. The most important features characterizing the rule are coverage and accuracy. Coverage describes whether number of
recognized objects is good enough, and accuracy determines how much we can trust the rule –
the rule should correctly classify the most of recognized objects. Rules matching both
conditions are considered to be high quality rules. We expect these rules to represent general
regularities found in data.

Another important criterion used for created rules estimation is the number of descriptors.
Generally speaking we are interested in rules with the minimal number of descriptors.

Measures used to estimate the rule are called rule quality measures. Two most important
measures are coverage and accuracy:

\[ q_{\text{coverage}} = \frac{n_{e\psi}}{n_{\psi}} \quad \text{and} \quad q_{\text{accuracy}} = \frac{n_{e\psi}}{n_{\psi}} \]

where \( n_{e\psi} \) denotes the number of object recognized by the rule \( \varphi \rightarrow \psi \), \( n_{\psi} \) denotes the number
of objects that belong to the decision class described by the rule \( \varphi \rightarrow \psi \) and \( n_{e\psi} \) denotes the
number of objects supported by the rule.

There are many different rule quality measures and they usually try to combine coverage and
accuracy measures. More detailed description of other measures can be found in [7][8].

## 5 MODLEM algorithm

Conditional attributes can be attributes of different types. Depending on the attribute type,
rangle of attribute value and descriptor form can vary.

Let \( a \in V_a \) be a single descriptor of the rule \( \varphi \rightarrow \psi \). Particularly if the attribute \( a \) is a symbolic
attribute, the set \( V_a \) is a single element set. In such case we use notation \( a = v_a \). If attribute \( a \) is
numeric attribute the range of \( V_a \) values may have one of the following forms: \((-\infty, v_a \),
\([v_a, +\infty), [v_{a1}, v_{a2}]\), where \( v_a, v_{a1}, v_{a2} \in V_a \).

The MODLEM algorithm was introduced by M. Stefanowski [4]. This heuristic algorithm
creates minimal set of rules which covers all decision classes form a training set. Output of
the algorithm is set of strong decision rules generated from numeric data without its prior
discretization.

Short description of how the algorithm works is presented below. The pseudo code of the
algorithm may be found in [4].

The algorithm starts with the whole set of objects – \( \text{Uncovered} \) set, and the empty set of
decision rules. During the process objects covered by induced rules are removed form the
\( \text{Uncovered} \) set. Method of the rule induction is similar to the construction of a decision tree.

For each conditional attribute from the \( \text{Uncovered} \) set, all of its values form the current
\( \text{Uncovered} \) set are being tested in order to find the best limit point \( g_a \).

The limit point \( g_a \) is the value placed between two successive values of the attribute \( a \) (for
example \( v_a < g_a < w_a \)) and divides the current range of the attribute \( a \) values into two subsets –
values that are bigger than \( g_a \) and smaller than \( g_a \). This limit point \( g_a \) splits also the current set
of training objects - \( \text{Uncovered} \) into two subsets \( U_1 \) and \( U_2 \). Only values between attributes
from different decision classes are considered as possible limit points.

Optimal limit point is the point which minimizes the value of the following expression:

\[
\frac{|U_1|}{|U_0|} + \text{Entr}(U_1) + \frac{|U_2|}{|U_0|} \text{Entr}(U_2)
\]

where \( \text{Entr}(U) \) stands for entropy of the set \( U \).

One of two intervals is chosen as the conditional descriptor. That is the one from adequate
sets \( U_1 \) and \( U_2 \) which includes more objects of considered decision class. Chosen descriptor is
added to descriptors created in previous steps of the algorithm making together with them (in
the form of the conjunction of conditions) the conditional part of the rule.
The rule creation process stops when all objects form Uncovered set are included in the decision class described by the rule (when all objects are covered by that rule). In other words, rules created by the MODLEM algorithm are accurate (the bottom approximation) or accurate “as much as it is possible for the training set” (the top approximation) [6]. The algorithm creates coverage of a given decision class, then – after the rule generation – all objects supporting that rule are removed from the training set and the algorithm is processed on the remaining set of objects.

6 Experimental results

It is very difficult to define one universal criterion which can be used to compare performance of the rule induction algorithms. For example different criterion is used when the rule set is generated for the classification purpose and different, when obtained rules are used to discover some previously unknown patterns in data.

As presented algorithms generate rules for further classification, two following criterions were used for algorithms estimation: the number of generated rules and their classification accuracy (3).

To compare algorithms performance three benchmark data sets form the UCI Repository of machine learning databases [13] where used:

- breast: 9 conditional attributes, 2 decision classes, 286 objects
- lympho: 18 conditional attributes, 4 decision classes, 134 objects,
- diabetes: 7 conditional attributes, 2 decision classes, 758 objects.

The TRS library, in which one can find implementation of all presented algorithms, was used to compare algorithms performance. The library is a tool for inducing and postprocessing decision rules, created and developed on the Silesian University of Technology [14].

For all data sets the 10-fold cross validation testing method was used. Results of experiments are presented in the table below.

Table 1  Comparison of induction rules algorithms

<table>
<thead>
<tr>
<th></th>
<th>Breast</th>
<th>Diabetes</th>
<th>Lympho</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of rules</td>
<td>Classification accuracy</td>
<td>Number of rules</td>
</tr>
<tr>
<td>Reduct based algorithm</td>
<td>1399</td>
<td>0.727273</td>
<td>1560</td>
</tr>
<tr>
<td>Sequential covering algorithm</td>
<td>490</td>
<td>0.69697</td>
<td>821</td>
</tr>
<tr>
<td>MODLEM algorithm</td>
<td>104</td>
<td>0.575758</td>
<td>244</td>
</tr>
</tbody>
</table>

7 CONCLUSIONS

This paper presents several approaches to the decision rules induction problem. The first presented algorithm generates exhaustive set of rules, the second and the third ones are heuristic algorithms generating minimal set of rules form the coverage. The MODLEM algorithm also allows generating rules from numeric data without its prior discretization. Several simple rule quality measures and its application were also introduced. Results of the algorithms comparison with regard to number of generated rule sets and rules classification abilities show that the MODLEM algorithm generates much smaller number of rules than other algorithms, while its classification accuracy remains almost as good as the accuracy of other algorithms. In all cases the best result of the classification accuracy is achieved by the reduct based algorithm but the number of generated rules is always
significantly higher than the number of rules obtained by the sequential covering or MODLEM algorithm. For practical reasons it is better to select algorithm generating less rules as long as it classification accuracy is not significantly lower. Almost all of induction rules algorithms used in rough sets theory face the same problem related to the huge amount of data which usually must be processed. Decision tables most often consist of a dozen or so attributes and several thousand of objects. For so large decision tables rule induction algorithms in their standard versions can not be performed fast enough. Solution for that could be to adapt existing algorithms for distributed or parallel environment. Further investigation of the author of this article is to propose parallel extensions for known decision rules algorithms.

8 References

Joint trade of flowgate rights

Przemysław Kacprzak
Warsaw University of Technology, Faculty of Electronic and Information Technology
P.Kacprzak@elka.pw.edu.pl

In this paper method for decentralized trading in transmission rights is proposed. This method is based on flowgate rights and multi-commodity exchange. Customer needed point-to-point rights are composed of many flowgate rights. Because multilateral are more effective joint exchange of energy, options and flowgate rights is proposed. Column generation technique is proposed to speed calculation.

1 Introduction

Electric energy companies started as vertically integrated enterprises. Then came liberalization of energy markets. Electric companies had to provide access to their network for other parties on non-discriminative rules. At first contract path approach was used. Example is shown on figure 1. When client in node A wanted to buy energy from supplier in node E he had to buy transmission rights for path from E to A – shown in red. But real power flow was different – as show in blue. Most important proposed workarounds of this problem were flowgate rights (FGR) and coordinated auctions.

In case of coordinated auctions regional transmission organization (RTO) receives all offers for point-to-point rights and assigns them. The problem is limited liquidity of this market – not many other market participants are interested in the same

In case of FRGs RTO at first has to determine existing constraints – flowgates - in system and impact of possible power flows, by publishing power transfer distribution function (PTDF) tables. Then consumers willing to use transmission grid has to buy appropriate FGRs. Market of FGRs can be much more liquid than point-to-point rights market. Oponents argue [2] that first there are too many flowgates for most customers to cope (not many is active but many could be active), second resulting schedules won't meet security requirments.

First problem can be solved by using multi-commodity exchanges instead of single-commodity markets. Second problem can be solved by bringing in more flowgates.

2 Multi-commodity exchange framework

In the traditional single-commodity exchange market operator calculates such a clearing price, that all buy offers with higher prices are accepted and and all sale offers with lower prices are
accepted. All offerents pay (or are paid) clearing price. This approach has good properties, such as maximizing social welfare. It could be modeled as following linear programming problem:

\[
\begin{align*}
\text{max} & \quad \sum_{m \in B} e_m d_m - \sum_{l \in S} s_l p_l \\
\sum_{m \in B} d_m & \leq \sum_{l \in S} p_l \\
0 & \leq p_l \leq p_{l,\text{max}} \quad l \in S \\
0 & \leq d_m \leq d_{m,\text{max}} \quad m \in B
\end{align*}
\]

Where \( B \) is a set of buy offers, \( S \) is a set of sale offers, \( e_m \) is price of \( m \) buy offer, \( s_l \) is price of \( l \) sale offer, \( d_{m,\text{max}} \) and \( s_{l,\text{max}} \) are amounts of demanded/offered goods. Objective function 1 is responsible for choosing buy offers with highest price and sale offers with lowest price, constraint 2 is balance constraint, constraints 3 and 4 are boundaries resulting from offered volume. The clearing price is taken as dual value of constraint 2.

Unfortunately this approach doesn’t allow to jointly trade several commodities, e.g. it’s impossible to guarantee that offers from given subset of offers will be all accepted or all rejected. It’s possible to enhance this model as in [1] to allow this. Then offers can contain more than one commodity. Offers containing single commodity Offers containing more than one commodity bought in the constant proportion are called integrated buy offers. Proportions are determined by commodity indicators \( a_i \). If commodity A has a indicator 1 and commodity B has indicator 0.5 then for each unit of commodity A half unit of commodity B will be bought. Market price of integrated offer is evaluated as sum of prices of contained commodities weighted by their indicators.

This could be modeled as following linear programming problem. There are \( n \) commodities. \( a_{l,i} \) is \( i \) commodity indicator in \( l \) sale offer, similarly \( a_{m,i} \) is \( i \) commodity indicator in \( m \) buy offer.

\[
\begin{align*}
\text{max} & \quad \sum_{m \in B} e_m d_m - \sum_{l \in S} s_l p_l \\
\sum_{m \in B} a_{m,i} d_m & \leq \sum_{l \in S} a_{l,i} p_l \quad i = 1..n \\
0 & \leq d_m \leq d_{m,\text{max}} \quad m \in B \\
0 & \leq p_l \leq p_{l,\text{max}} \quad l \in S
\end{align*}
\]

This model looks like previous one. Objective function 5 is the same as 1, boundaries 7 and 8 are the same as 3 and 4. Balance constraint 6 is different, it’s not just one constraint, it’s a
group of constraints – one for each commodity. Market prices of commodities are taken as dual values of corresponding balance constraints.

3 Transmission offers

3.1 Principles
For each line there are two flowgates – each for one direction. Customers need point-to-point transmission rights, which could be obligations or options. If transmission is certain to happen the it can create counterflow, which relieves some constraint. In case of obligation transmission rights customer has to buy FGRs in amounts indicated by PTDFs – if PTDF is negative then in fact he sells the FGRs. In case of option transmission rights customer buys FGRs if PTDF are positive. Owners of options are not sure to produce counterflows, so they can't sell flowgate rights. Selling flowgate rights is connected with exercising option.

3.2 Example

On figure 2 simple network is shown. All lines has the same admitance. PTDFs are given in Table 1.

Table 1 PTDF table of network on fig. 2

<table>
<thead>
<tr>
<th>Line</th>
<th>Flow</th>
<th>AB</th>
<th>AC</th>
<th>BC</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>0.667</td>
<td>0.333</td>
<td>-0.333</td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>-0.333</td>
<td>0.333</td>
<td>0.667</td>
<td></td>
</tr>
<tr>
<td>L3</td>
<td>-0.333</td>
<td>-0.667</td>
<td>-0.333</td>
<td></td>
</tr>
</tbody>
</table>

There would be 6 flowgates: L1+,L1-,L2+,L2-,L3+,L3-. So in order to get transmission right customer has to buy amounts of FGRs as shown in Table 2 and Table 3.

Table 2 Resources for obligation transmission rights

<table>
<thead>
<tr>
<th></th>
<th>AB</th>
<th>BA</th>
<th>AC</th>
<th>CA</th>
<th>BC</th>
<th>CB</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1+</td>
<td>0.667</td>
<td>-0.667</td>
<td>0.333</td>
<td>-0.333</td>
<td>-0.333</td>
<td>0.333</td>
</tr>
<tr>
<td>L1-</td>
<td>-0.667</td>
<td>0.667</td>
<td>-0.333</td>
<td>0.333</td>
<td>0.333</td>
<td>-0.333</td>
</tr>
<tr>
<td>L2+</td>
<td>-0.333</td>
<td>0.333</td>
<td>0.667</td>
<td>-0.333</td>
<td>0.667</td>
<td>-0.667</td>
</tr>
<tr>
<td>L2-</td>
<td>0.333</td>
<td>-0.333</td>
<td>-0.333</td>
<td>0.333</td>
<td>-0.667</td>
<td>0.667</td>
</tr>
<tr>
<td>L3+</td>
<td>-0.333</td>
<td>0.333</td>
<td>-0.667</td>
<td>0.667</td>
<td>-0.333</td>
<td>0.333</td>
</tr>
<tr>
<td>L3-</td>
<td>0.333</td>
<td>-0.333</td>
<td>0.667</td>
<td>-0.667</td>
<td>0.333</td>
<td>-0.333</td>
</tr>
</tbody>
</table>
Table 3 Resources for options transmission rights

<table>
<thead>
<tr>
<th></th>
<th>AB</th>
<th>BA</th>
<th>AC</th>
<th>CA</th>
<th>BC</th>
<th>CB</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1+</td>
<td>0.667</td>
<td>0</td>
<td>0.333</td>
<td>0</td>
<td>0</td>
<td>0.333</td>
</tr>
<tr>
<td>L1-</td>
<td>0</td>
<td>0.667</td>
<td>0</td>
<td>0.333</td>
<td>0.333</td>
<td>0</td>
</tr>
<tr>
<td>L2+</td>
<td>0</td>
<td>0.333</td>
<td>0.333</td>
<td>0</td>
<td>0.667</td>
<td>0</td>
</tr>
<tr>
<td>L2-</td>
<td>0.333</td>
<td>0</td>
<td>0</td>
<td>0.333</td>
<td>0</td>
<td>0.667</td>
</tr>
<tr>
<td>L3+</td>
<td>0</td>
<td>0.333</td>
<td>0</td>
<td>0.667</td>
<td>0</td>
<td>0.333</td>
</tr>
<tr>
<td>L3-</td>
<td>0.333</td>
<td>0</td>
<td>0.667</td>
<td>0</td>
<td>0.333</td>
<td>0</td>
</tr>
</tbody>
</table>

Let all lines has capacity 100 MW. Customers submit following offers as in Table 4.

Table 4 Offers for transmission

<table>
<thead>
<tr>
<th>No</th>
<th>Source</th>
<th>Sink</th>
<th>Type</th>
<th>Price</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>B</td>
<td>Obligation</td>
<td>10</td>
<td>180</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>C</td>
<td>Obligation</td>
<td>10</td>
<td>120</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>A</td>
<td>Obligation</td>
<td>-5</td>
<td>150</td>
</tr>
<tr>
<td>4</td>
<td>B</td>
<td>C</td>
<td>Option</td>
<td>11</td>
<td>90</td>
</tr>
<tr>
<td>5</td>
<td>A</td>
<td>C</td>
<td>Option</td>
<td>12</td>
<td>180</td>
</tr>
</tbody>
</table>

In this example client offered negative price for transmission right. He is sure to paid if his offer is accepted. His offer will be accepted if it really helps to relieve some constraint.

In this example FGR L3’ has price of 18, price of the other FGRs are 0. Offers 1,2,3,4 are accepted in whole, offer 5 is accepted partially. Market price of offers 1,2 and 4 is 6, of offer 3 is -12, and offer 5 is 12.

4 Joint exchange of energy and transmission rights

The joint energy, options and transmission rights exchange would be more useful than just transmission rights exchange. The basic commodities would be energy in different nodes, call and put options in different nodes and FGRs. Markets operator has to use available FGRs in a
way that maximizes social welfare function. It would be fairly simple if exclude options. In addition to customer offers market operator has to submit offers of energy transmission, which consists of buying energy in one node, selling in the other node and buying and selling FGRs. Spanning tree of these is enough to get the best result. Transmission between other nodes can be set as combination of base transmission offers.

For transmission of options even all node-to-node transmission offers are not enough. First option transmission combined of base transmission offers possibly use more more resources. Example of this is shown on figure 3. In the gray area are possible flows if client buys transmission rights From A to B and from B to C, on vector AC are possible flows if client buys just transmission right from A to C. Second some flowgate requirements can be mitigated with properly allocated energy options. To simplify let's assume that only line L1 in the network shown on fig 2 can be active constraint. Options are supplied in nodes A and B, in node C options are bought. Because L1 has capacity of 100 MW, from each node 300 MWh options can be transferred (if basic period is one hour). But if generation is increased simultaneously in nodes A and B the power flow in line L1 is constant.

Best allocation of network capacity for optional transmission is difficult task. One has to evaluate all $2^n$ scenarios of options exercising. It was show in [3] that this problem can be relaxed by choosing proper subset of scenarios. Different approach was used in experiments. Column generation method was used to make offers of cheapest options acquisition. There was used following subproblem:

$$
\text{max } Q = \pi^+ p^+ + \pi^- p^- + \sum_{f \in F} \pi_f u_f - \sum_{n \in N} \pi^+_n d^+_n - \sum_{n \in N} \pi^-_n d^-_n
$$

$$
u_f \geq \sum_{n \in N} PTDF(f,n)(a^+_n - a^-_n) \quad \forall f \in F
$$

$$
u_f \geq \sum_{n \in N} PTDF(f,n)(b^+_n - b^-_n) \quad \forall f \in F
$$

$$\quad a^+_n \leq d^+_n, b^+_n \leq d^+_n, a^-_n \leq d^-_n, b^-_n \leq d^-_n \quad \forall n \in N
$$

$$\quad p^+ = \sum_{n \in N} a^+_n - \sum_{n \in N} a^-_n
$$

$$\quad p^- = \sum_{n \in N} b^+_n - \sum_{n \in N} b^-_n
$$

$$\quad 0 \leq p^+, p^- \leq 1
$$

Fig 3: Possible flow with option transmission rights
This subproblem is solved for each node where is demand for options. \( \pi^+ \) is price of call option in designated node, \( \pi^- \) is price of put option in designated node, \( N \) is set nodes where sale offer of options are set, \( F \) is a set FGRs, \( \pi^+_n \) is price of call option in node \( n \), \( \pi^-_n \) is price of put option in node \( n \), \( a^+_n \) and \( a^-_n \) are number of call and put options exercised when call option in designated node is exercised, \( b^+_n \) and \( b^-_n \) are number number of call and put options exercised when put option is exercised in designated node. Objective function is responsible for choosing cheapest way of providing options. Constraints 10 and 11 are responsible for evaluating amount of needed FGRs. Constraints 12 ensure that it won’t be exercised more options than there were bought. Constraints 13 and 14 ensure that enough options are exercised to provide requied energy.

5 Numerical experiments

Different methods were examined for efficiency and speed. The greater welfare function (w.f) the more efficient method. Optimal method gives best possible welfare function. In simple method all possible node-to-node options transmission offers are used. In column generation (c.g.) additional offers are generated (from many nodes to one node). The calculation were made on computer PIV 3GHz with 512 MB of RAM using package CPLEX 9.1.

Table 5 Numerical experiments

<table>
<thead>
<tr>
<th>Nodes</th>
<th>Nodes with option demand</th>
<th>Lines</th>
<th>Optimal time[s]</th>
<th>Optimal w.f.</th>
<th>Simple time[s]</th>
<th>Simple w.f.</th>
<th>C.g. time[s]</th>
<th>C.g. w.f.</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>3</td>
<td>6</td>
<td>~0</td>
<td>1368</td>
<td>~0</td>
<td>1288</td>
<td>~0</td>
<td>1368</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>44</td>
<td>0.5</td>
<td>49526</td>
<td>0.02</td>
<td>49444</td>
<td>0.03</td>
<td>49444</td>
</tr>
<tr>
<td>16</td>
<td>12</td>
<td>24</td>
<td>613</td>
<td>29836</td>
<td>0.02</td>
<td>28890</td>
<td>0.14</td>
<td>29762</td>
</tr>
<tr>
<td>30</td>
<td>12</td>
<td>70</td>
<td>2427</td>
<td>83243</td>
<td>0.16</td>
<td>82427</td>
<td>1.24</td>
<td>82883</td>
</tr>
</tbody>
</table>

6 Conclusions

Presented method can used to implement decentralized energy market. Multi-commodity exchange is useful for multilateral trade, which is more effective than just bilateral trade. Proposed method of generating additional transmission of energy options offers doesn’t give best possible results. On the other hand it’s quick, can easily be computed in parallely, so it may be preferred in continuos trading.

7 References

Abstract: This paper attempts to explain an image compression technique based on fuzzy-wavelets. This technique combines the multiresolution features of wavelet analysis and a fuzzy system, which complements the wavelet decomposition. Fuzzy system is constructed with the coarsest scale at first and a combination with wavelet multiresolution system is coming next. Introduced type of top-down scale decomposition is practical in progressive image transmission with a high quality, which can be used in medical applications.

1 Introduction

The wavelet transform is a mathematical transform similar to the commonly known Fourier transform. Wavelet analysis is a form of “multiresolution analysis”, which means that wavelet coefficients for a certain function contain both frequency- and time-domain information. This fact makes wavelets useful for signal processing applications where knowledge of both frequency information and the location in time of that frequency information is useful. The usefulness of wavelets in image compression and transmission lies in the fact that the wavelet transform clearly separates high-pass and low-pass information on a pixel-by-pixel basis. If high compression and fast transmission rates are desired, only the low pass coefficients need to be dealt with. However, if bandwidth allows for higher quality transmission, some high pass coefficients may be sent to sharpen the image quality. Many of wavelet image compression algorithms use a technique called subband coding.

The space-frequency localization of subband image data takes advantage of efficient data structures for spatial correlation. Generally, correlation is used for the signal similarity detection, and because WT use the input signal correlation with a variously dilated mother wavelet, this transform detects the input signal similarity to variously expanded mother wavelet versions.

Wavelet coefficients have the property that related coefficients in different scales are located in a same orientation and location as the wavelet hierarchy. Except the subbands with the highest resolution, every coefficient in a given scale is related to a set of coefficients in the next finer scale orientation and location. The mother wavelet, in a consequence of a zero mean value, has always behaviour, which represents slower or faster change. That is why wavelet transform is used to find such changes position in an input signal.

This feature of WT is also very useful in an image processing. Medical images are well suited to wavelet processing as often the information is contained in sharp edges or localized contrasts. Contours and edges of the object in an image are very important for a human being to resolve what image actually represents. On the basis of contours we distinguish the shapes and individual image elements. These contours and edges are represented in images by a significant color or brightness change, which is easily detectable by wavelet transform.
Wavelet transform shows high signal values in a place of noticeable edges, and values of zero in flat areas with constant intensity or color. This means, that very few pixels are required in the smooth regions of an image. Preserving these high values corresponding to the image edges and removing all the others allows a significant data reduction and to preserve the basic image structure at the same time. This considerable data reduction leads in its consequence to high compression rates with a good image quality.

The wavelet transform is not utilizable only in data compression, but also in a high-quality image transmission in slower networks as Internet and so on. The principle of such a transmission lies in a transfer of a small amount of wavelet coefficients at first, from which the rough image version is reconstructed. In a time progress, other coefficients are added to complete the blurry image details. But the fast wavelet decomposition algorithm cannot be applied at extremely fine scales. A fuzzy-wavelet system is advantageous to use in such a case. In the next part we formulate the algorithm of forming such a system.

2 Image compression

An adaptive fuzzy system is constructed at coarse scale at first, so the first system version contains only rough characteristics of the original image. Fuzzy rules and membership functions can be than extracted with a use of the least-squares fitting or other methods [2]. We use a specific type of wavelets known as B-spline wavelets. Because the fuzzy system allows us to combine linguistic knowledge with a numerical one, the solution becomes more synoptic. If we than use a wavelet basis of multiresolution, we can set-up a wavelet system deduced from the residue, which improves the operation by accessing fine scales. The residue is transferred to the next finer scale, to repeat the wavelet decomposition until it becomes smaller than predefined error of compression. We have to perform a redundant fuzzy rules and wavelet coefficients reduction, to increase the compression ratio. This algorithm, as mentioned above, leads to a kind of pyramidal structure, which is advantageous in image transmission. Although the number of fuzzy rules and wavelet coefficients in coarse scales is quite small, the reconstructed image still includes enough information of the original one. We need to quantize the fuzzy rules and wavelet coefficients before the process of encoding. For this purpose we use a hard quantization method. The encoding itself can be carried out by two different techniques. Because in coarse scales most of the fuzzy rules and wavelet coefficients do not have values of zero, they can be transmitted in a given exact order. Before the transmission itself, the entropy encoding has to be applied to the rules and coefficients. A good encoder, generally, takes advantage of the wavelet coefficients structure as a multiscale and spatial non-correlation.

At first, we focus on a fuzzy-wavelet system in itself. At first we introduce a specific kind of a wavelet scaling function: a B-spline.

A B-spline function is a piecewise polynomial function defined on a lattice [1]. The spline function order determines the properties of the spline. The simplest spline function is the characteristic function of the unit interval. The characteristic function $N^1(x)$ or Haar function (Haar, 1910) is defined as:

$$N^1(x) = \begin{cases} 1 & \text{if } 0 \leq x < 1 \\ 0 & \text{otherwise} \end{cases}$$

The second order cardinal B-spline is a triangular function, a continuous function summing piecewise polynomials of order 1.
Cardinal B-spline $N^k(x)$ of order $k$ can be generated by repeated convolution of the characteristic function:
\[
N^k(x) = \int_0^x N^{k-1}(x-t)dt \quad \text{or} \quad N^k = N^1 * N^{k-1}
\] (2)

B-splines have a number of important and useful properties:
- The B-spline functions are the polynomial splines with the shortest support.
- All values are positive.
- Splines have a closed-form formula as their piecewise polynomials.
- Spline functions can be used to form a partition of unity, by using a superposition of translated B-splines or:
\[
\sum_{j=0}^{\infty} N^k(x+j) = 1
\] (3)

This property is very useful in order to give a fuzzy interpretation to a spline decomposition.

Also biorthogonal spline-wavelets have a number of properties that are quite useful in real applications. On one hand, both the wavelet $\Psi(x)$ and its dual $\tilde{\Psi}(x)$ have a compact support, on the other band, the scaling function $\phi(x)$ is always positive. This permits to interpret the scaling functions as membership functions in a fuzzy framework. Biorthogonal spline-wavelets are typically used in wavelet networks and also in on-line problems in which a simple method to process the boundaries is necessary.

As mentioned above, we choose the coarsest scale at the beginning ($j = j_0$) for the fuzzy system constitution [2]. Because the image is defined on a bounded unit square $[0, 1]^2$, there are instituted both inner B-spline fuzzy membership functions and boundary membership functions. The boundary membership functions and wavelets institution has a reason of avoiding the products of a compressed image. General method to handle the boundary effect is to use the unique form of fuzzy membership function and wavelet, but extend the image beyond $[0, 1]^2$, which, nevertheless, result in discontinuities in boundary area anyway, and more fuzzy rules and wavelets have to be used. To make sure that there is at least one inner membership function, the coarsest scale must fit the condition $2^{j_0} \geq 2k+1$, where $k$ is an order of a cardinal B-spline. To extract the fuzzy rules and coefficients, a compact number of sample pixels is needed. Generally speaking, we need to shift the given pixel set to fit the single scales, which can be done with a use of image averaging filter. This is a simple task of dividing the image into $2^j \times 2^j$ squares with a side of $2^{n-j}$, where $j = j_0$ at start. After this process we’ll have a sample pixel set $S$ at the scale $j$, as denotes (4), where ordered pairs $(x_i, y_i)$ are the each square centers.

\[
S_j = \left\{(x_i, y_i) \in Z \mid \text{average}(z_i, S_i)\right\}
\] (4)

Then we perform the fuzzy quantization on our coarsest pixel set $S_{j_0}$. Fuzzy quantization is a generalization of a hard quantization schema and the best known and most widely used fuzzy quantization technique is the fuzzy c-means algorithm developed by Dunn and refined by Bezdek. The objective of FCM segmentation is to compute the cluster centers and generate the class membership matrix. To quantize the sampled image, we must choose the right number of fuzzy quantities. For the coarsest scale this number is equal to
\[
L_F^{j_0} = (2^{j_0} + k) \times (2^{j_0} + k),
\] where $k$ is a cardinal B-spline order.
The fuzzy quantization $z(i_1;i_2) = f(i_1 \cdot 2^{-n}; i_2 \cdot 2^{-n})$ where $i_1; i_2 \in [0;2^n - 1] \cap Z$, is performed by the unit square $[0;1]^2$ segmentation into a square grid of $\delta_j = \frac{1}{2^j}$ (this formulation refers to square images with a size of $2^n \times 2^n$). Let us label the $i$th point $i = (i_1; i_2)$ coordinate as $p_i$, which is the center of the $i$th membership function $\mu_{i_1,i_2}(x,y)$. A fuzzy quantization on $[0;1]^2$ is then defined as a set of such membership functions.

Reversely, to map the center grade pairs to a real number, we have to perform a defuzzification process. If the center grade pairs are defined as $\{q_i, \lambda_i\}$, then the defuzzification is defined as $z = \sum q_i \lambda_i$.

After the defuzzification process we obtain an equation $z = \mu_i(x,y)G_F(p_i) = \sum q_i \mu_i(x,y)$, where $G_F(X)$ is a set of fuzzy rules. Let us introduce the expression $Z = G_F(X)$ for the fuzzy rule set description, where we designate X and Z as fuzzy variables. A fuzzy variable can be defined as a variable that takes on fuzzy quantities within the framework of its definition region. Fuzzy rules are to operate with the variables and to provide the fuzzy system’s performance.

After the quantization is finished, the partial compressed image has the form of:

$$I_{j_0} = Q_{j_0} \Lambda_{j_0}$$

(5)

$I_{j_0} = [z_{j_0}^0, \ldots, z_{L_{j_0}}^0]$ is a vector $Q_{j_0}$ is a 1x$L$ fuzzy rule vector and matrix $\Lambda_{j_0} = [\mu_{i_1}^0(x_{i_1}^0, y_{i_1}^0), \ldots, \mu_{i_L}^0(x_{i_L}^0, y_{i_L}^0)]$, with $i = 1, 2, \ldots, L_{j_0}^0$.

There are many methods to extract the fuzzy rules with. After their extraction we’ll get a partial compressed image as shown in (6):

$$I_{j_0}^0 : z_i^0(x,y) = \sum q_{j_0,i} \mu_i^0(x,y)$$

(6)

From the other hand, the wavelet decomposition is also needed to complete the scheme. The univariate B-spline scaling function and mother wavelet on the bounded interval $[0, 1]$ are defined as $\phi_{b,k}^j(x) = N_{b,k}^j(x)$ where $b = -k, -k+1, \ldots, 2^j - 1$.

Splines are commonly used in function approximation, and multivariate B-spline basis functions can be formed by multiplying univariate basis functions. This is exactly the way of our approach in representing the image intensity function. For the two-dimensional intensity function, we need to form a basis consisting of three functions. The set $Ψ$ is therefore defined as $Ψ = \{\phi(x)\psi(x); \psi(x)\phi(y); \psi(x)\psi(y)\}$, where $\phi$ is a univariate scaling function and $\psi$ its corresponding wavelet. This representation can be also called the mutiresolution analysis.

After we defined the multiresolution and we can represent the image a multiscale fuzzy-wavelet as (7).

$$I \sim z(x; y) = \sum q_{j_0,i} \mu_i^0(x,y) = \sum q_{j_0,i} \mu_i^0(x,y) + \sum_{j=0}^{j_1} \sum_{d \in Z^2} \sum_{\varphi \in \Psi} c_{j,d} \varphi_{j,d}(x; y)$$

(7)

Now we can start with the wavelet system construction. At the beginning we also choose the coarsest scale, and than perform the wavelet decomposition as $R_{j_0} = C_{j_0} \Psi_{j_0}$, where $R_{j_0} = [r_{j_0}^0, \ldots, r_{L_{j_0}}^0]$ is a residual vector using image averaging on the sample pixel set $S$, $C_{j_0}$ is a vector of wavelet coefficients, and $\Psi_{j_0} = [\varphi_{j_0,b}^0(x_{i_1}^0, y_{i_1}^0), \ldots, \psi_{j_0,b}^0(x_{j_{L_{j_0}}}^0, y_{j_{L_{j_0}}}^0)]$, $\varphi \in \Psi$. 

154
is a scaling function matrix. After obtaining the wavelet coefficients we can describe the image as:

$$I_{j_0} : z^{j_0}(x, y) = I_{j_0}^0 + \sum_{b \in Z} \sum_{b' \in \Psi} c_{j_0,b} \mathbf{u}_{j_0,b}(x, y)$$  \hspace{1cm} (8)

Now we continue the process by calculating the residue, which is the difference between $I$ and $I_{j_0}$ on the sample pixel set $S$. With the residue we perform the same operations in the next cycle where we move $j = j + 1$.

This process of decomposition must be recursively repeated at all following scales, until the residue falls below the given compression error. In general, the wavelet coefficients must be updated until $\parallel C_{j,i+1} - C_{j,i} \parallel < \frac{\varepsilon}{2}$, or the improvement between two cycles becomes insignificant.

After the coefficients are updated at each scale, a thresholding process must be performed. This means, that we have to get rid of those coefficients, which are insignificant at the compression accuracy issue. The thresholding condition is defined as in (9):

$$\parallel C_{j,i} \parallel < 2^{-j} \cdot \varepsilon$$  \hspace{1cm} (9)

The condition (9) is applied to each coefficient, and the ones, that don’t meet the criterion, are removed from the system. This operation encloses the process of the system construction. The last step is the resulting data encoding.

An entropy encoding is a coding scheme that assigns codes to symbols so as to match code lengths with the probabilities of the symbols. Entropy encoders are used to compress data by replacing symbols represented by equal-length codes with symbols represented by codes proportional to the negative logarithm of the probability. This results in the fact, that the most common symbols use the shortest codes. The most common entropy encoding techniques are Huffman coding, range encoding, and arithmetic coding.

### 3 Conclusion

Multiresolution analysis is an expedient method to discover important signal features. It also permits to compare features at different resolutions. Multiresolution permits to identify features such as transitions, edges and peaks, which is useful in many problems’ solution. There are some methods combining fuzzy logic and multiresolution analysis, developed for the purpose of segmentation and contour extraction. Sample pixels should contain, for the most part, information about edges, only few pixels are needed in smooth areas. Introduced type of top-down scale decomposition is practical in progressive image transmission with a high quality, which can be used in medical applications.

### 4 References

Using Genetic Algorithm for Fractional Systems Identification

Pavel Petránek¹, Abdelbaki Djouambi² and Alina Besançon-Voda ²
¹ VSB – TU Ostrava, Department of Measurement and Control, Czech Republic
² Laboratoire d’Automatique de Grenoble, France
pavel.petranek@vsb.cz, djouambi@lag.ensieg.inpg.fr, alina.voda@inpg.fr

Abstract: Genetic algorithms are widely used universal optimization procedures in many different areas. Their great advantage is evolutionary access for solving problems. Because of this feature they are used for problems which are very difficult to solve by classical mathematical methods or using of this classical approach is not possible. This paper considers an evolutionary approach for fractional system identification and attempts to show how genetic algorithms can be applied in fractional systems identification tasks. Some study cases confirm that good performance can be achieved by this method.

1 Introduction

One of the first steps in many technological areas consists of building a mathematical model to represent dynamical behavior of systems [1]. So building a linear, or non linear model, is a common and often difficult problem.

It is possible to divide models into two groups [2]: phenomenological and behavioral models. In the first group, parameters have a meaning, which can be useful to validate the model. Behavioral models try to approximate process evolution without prior information. Studies on real systems such as thermal [3], or electrochemical [4], reveal inherent non integer (or fractional) differentiation in their behaviour. The use of the integer models is thus inappropriate in identifying these non integer systems.

There is a well established set of identification techniques for linear integer order models, but identification is not so clear with non-integer models. In this case parameter identification becomes a complicated optimization problem. For these situations, a global optimization technique is necessary, and Genetic Algorithms offer a very good approach to resolving this problem. In the following paper we will discuss specifics of the identification of parameters with genetic algorithm for a chosen structure of implicit fractional systems model. We use for validation comparisons of time and frequency responses of the identified and the corresponding real systems.

2 Fractional Implicit Model

Many of real physical systems can be represented by fractional order models [5] [6]. Most of them can be better represented by an implicit fractional transfer function given by the following Fractional Power Pole (FPP):

\[ F(s) = \frac{g_0}{(1 + \tau_0 s)^m} \]  \hspace{1cm} (1)

where
\[ g_0, \tau_0 \text{ and } m \in \mathbb{R}^+, 0 < m < 1 \]
This expression could be used only in case of simple transfer function. For more complicated fractional systems, they can be represented by a more general equation given as follow:

\[ F(s) = \prod_{i=1}^{p} \frac{g_i}{(1 + \tau_is)^{m_i}}, \]

where \( g_i, \tau_i \in \mathbb{R}^+ \) and \(-1 < m_i < 1\)

To understand the dynamic behavior of these fractal systems a full scale analysis has to be made. Because of their mathematical representation in the frequency domain are irrational functions, direct analysis methods and corresponding time domain behavior seems very difficult to handle. For the purpose of identification, analysis, synthesis and simulation of such irrational function, the need arises for a rational function approximation.

In order to represent the basic FPP transfer function given as (1) by a linear network, Y. Tsao and Onaral have proposed, see [5], a rational function approximation in the frequency domain which they called singularity function approach, as follows:

\[ F(s) \rightarrow F_{s}(s) = \prod_{k=1}^{N} \left( 1 + \frac{s^2}{z_k} \right) \]

where \( k = \frac{1}{\omega_0} \) from (1); \( p_k = \omega_0\sqrt{b} \); \( z_i = (ab)^{1-i} a p_i \),

\[ p_i = (ab)^{1-i} p_i, i = 1, 2, \ldots, N; \quad a = 10^{10^{1-\varepsilon}}, b = 10^{10^{1-\varepsilon}}, \quad ab = 10^{10^{1-\varepsilon}} \]

and \( \varepsilon \) represents approximation error \( a \) is defined as the location ratio of zero to a previous pole where \( b \) is defined as the location ratio of a pole to a previous zero, and \( ab \) the location ratio of a pole to a previous pole or zero to previous zero, \( m \) represents the slope of fractional system in Bode plot, it has been shown that.

\[ m = \frac{\log(a)}{\log(ab)} \]

The number of cells \( N \) is obtained when the frequency band \([\omega_0, \omega_{\text{max}}]\) is completely covered, given an error \( \varepsilon \), \( N \) can be given by the following equation:

\[ N = \text{Integer} \left( \frac{\log\left( \frac{\omega_{\text{max}}}{\omega_0} \right)}{\log(ab)} \right) + 1 \]

### 3 Genetic Algorithm

Genetic algorithms (GA) are adaptive search and optimization procedures derived from mechanisms of nature’s genetics [7]. The base for genetic operation is a string called chromosome, which codes the parameters of system. Chromosomes are built from genes. GA operates in cycles called generations. Copying and interchanging of chromosome’s parts reach transition between generations. Sum of all chromosomes in one generation is called population. Over generations, each member of the population is evaluated by fitness function and this value expresses how close this solution to the ideal chromosome is. GA’s mechanisms that produce next generation are crossover, mutation and reproduction.
Reproduction is a process that copies single chromosomes to new population according to their fitness value. Crossover transfers parts of selected chromosomes. Mutation changes with small probability values of genes in chromosomes. If we want to use GA to solve a problem, a few premises must be realized – code problem into appropriate chromosomes, find good fitness function and set termination criterion. For our example the chromosomes are based on combination of fractional system approximation. Each parameter is one gene of chromosome.

4 GA improvement

The greatest disadvantages of using GA are speed of convergence and possibility of getting stuck in local extreme. Our innovation approach is trying to increase algorithm speed. We speculate about standard duration and we found way how to enhance speed. Solution is in creation of expert module. Our expert module is inserted as layer inside a fitness function. When algorithm calls the fitness function the expert module is activated. The module has database of previous solutions (with appropriate fitness values) which is saved inside this module. There is possibility that chromosome wasn’t yet noticed. In such case the module lets fitness function to calculate fitness value and consequently it saves it for future use into database. Time needed to compute fitness function is the longest part of computation during complex problem solving.

5 Settings of GA

The system is classified by his response, which is caused by a pseudorandom binary input. The response is compared with model data. Quality of this response could be classified in many ways and will create base of fitness function. For our study we used extended standard equation for similar response validation \[8\]

\[
I = \Phi \cdot \left(101 - \frac{1 - \text{norm}(y - \bar{y})}{\text{norm}(y - \text{mean}(y))}\right) \cdot 100 \cdot 2^\Omega \tag{7}
\]

In the equation (7) \(\Phi\) is a symbol of penalization coefficient [9], which guarantees approximation bounds in chromosome. If one value is out of bound the final fitness value is doubled etc. For GA run is better to increase value of fitness function in case of unacceptable chromosome than simply discard such solution. There is possibility to get superior individual after crossover from two below the average chromosomes. The next symbol \(\Omega\) represents evolution pressure what prevents algorithm to get stuck in local extreme. It is applied only in case of positive result of this equation.

\[
\Omega = \sum (abs(y) - abs(\bar{y})) \tag{8}
\]

Initial population was picked randomly. Number of chromosomes in the population was six times larger than number of coefficients in chromosome. We used algorithm with two population and migration each 20 generations. Mutation was used Gaussian and for selection function was chosen stochastic uniform type. For termination condition was selected stall time with duration 500 s. Other parameters kept their default MATLAB settings.

6 GA Improvement Results

Expert module doesn’t consider only exact matches, but there is also zone of tolerance. Our module considers only fixed values of these differences (the speed of this module is very important for its efficiency).

Consider the model set given by the following equation,
As we verified really, small differences in the coefficients produced (Table 1) also only small differences in the output. The fit value [8] was computed from the following equation

\[ \text{fit} = \frac{1 - \text{norm}(y - \bar{y})}{\text{norm}(y - \text{mean}(y))} \times 100 \]  

In the table 1 is shown average difference in percent to the exactly the same curve for different number of changed coefficients.

**Table 1. Influence of coefficient differences to the fit value (9)**

<table>
<thead>
<tr>
<th>number of coefficient with difference 0,001</th>
<th>2</th>
<th>4</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>without noise (100-fit) [%]</td>
<td>0,02</td>
<td>0,08</td>
<td>0,11</td>
</tr>
<tr>
<td>with noise (100-fit) [%]</td>
<td>0,043</td>
<td>0,084</td>
<td>0,099</td>
</tr>
</tbody>
</table>

Our results are shown in tables below (Table 2). There are computation’s parameters (number of generations, average time needed for computation, identity of result with model and average time needed for computation of single generation) for runs with expert module and without this module.

**Table 2. Tables with results for computation with and without expert module**

<table>
<thead>
<tr>
<th>number of generation</th>
<th>expert module without noise</th>
<th>expert module with noise</th>
<th>without noise and without expert module</th>
<th>with noise and without expert module</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of generation</td>
<td>587,22</td>
<td>581,67</td>
<td>641,33</td>
<td>418,00</td>
</tr>
<tr>
<td>time of computation [s]</td>
<td>2177,31</td>
<td>2541,33</td>
<td>3418,76</td>
<td>2664,45</td>
</tr>
<tr>
<td>fit value [%]</td>
<td>99,49</td>
<td>98,47</td>
<td>99,36</td>
<td>98,56</td>
</tr>
<tr>
<td>time per generation [s/g]</td>
<td>3,71</td>
<td>4,37</td>
<td>5,33</td>
<td>6,37</td>
</tr>
</tbody>
</table>

We can clearly see benefit of this solution. The time reduction of computation is really remarkable. Average time of computation was decreased by 36.3% for system without noise and by 31.4% for noisy system.

**7 Numerical Simulations**

To illustrate performances of the proposed algorithm, we consider the fractional system which has the implicit form given by the following equation,

\[ F(s) = \frac{(1 + 10s)^{0.6}}{(1 + 3s)^{1.8}} \]  

\[ F(s) = \frac{g_0}{(1 + \tau_0 s)^m}, \quad 0 < m < 1 \]  

The problem is implemented by simulation using numerical transfer function given by the equation (12). The input is a PRBS.

\[ F_i(s) = \frac{1}{(1 + s)^{0.6}} \]  

We ran our identification process and results which we got are shown in (Table 3).
Table 3. Results of identification

<table>
<thead>
<tr>
<th></th>
<th>( g_0 ) (default)=1.0</th>
<th>( \tau_0 )</th>
<th>( m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>simulated system</td>
<td>1.000</td>
<td>0.600</td>
<td></td>
</tr>
<tr>
<td>identified without noise</td>
<td>1.003</td>
<td>0.599</td>
<td></td>
</tr>
<tr>
<td>identified with noise</td>
<td>0.991</td>
<td>0.609</td>
<td></td>
</tr>
</tbody>
</table>

For validation in time domain we used step signal. The result picture showed two curves which were almost overlapped. This result was identified from model without noise. The difference between models was not visible at that chart. Next step was validation of coefficient for system with noise. The noise is a white noise with maximum amplitude 20% of output signal. The same noise was used for all steps of identification. According to our expectations it can be seen that this solution approximate the searched parameters with an satisfactory accuracy.

Fig. 1 Difference of step output of identified system response (with noise – dashed line and without noise – full line) and etalon model

Fig. 2 Frequency responses of the identified system and the real system
a) without noise, b) with noise

We have proved also verification in frequency domain. The used frequency band was 10 decades up \( 10^1 \) to \( 10^9 \). Output charts show frequency response of etalon model (black line) and system without and with noise (Fig. 2). The system without noise (Fig. 2 – a) has very good frequency response because both curves are almost overlapping. The differences are very small. Frequency response of identified system with noise doesn’t look as similar in high frequency (Fig. 2 – b) as the displayed noiseless response.
8 Conclusion

This area is very interesting because of possible application into common technical problem. Our unorthodox access to solving this type of problems seems to be one of rare ways how to find suitable coefficient for implicit time domain model. The expert module which we implemented into GA is one of possible ways how to improve solution. Benefit of module is unexceptionable, but there are still features which we want to enhance speed, accuracy etc. The right direction we see in improving “intelligence” of the module primarily in way of precision and quickness. Our plan is to implement automatic investigation of precision dependence on the individual chromosome parameters and subsequently adaptation of their deposition. Another way of our work should be dedicated to improvement of possibilities of GA. There are still leaks in identification with sight to recognize order and coefficient of system. The complex identification is much more complex task than our example here. We trust that validation of our access is the first step in right direction. Right proof of our method we see in practical test. We would like to apply our method for data which will be available from medical biochip in short future. We expect that this test validates our method for practical use.

9 Acknowledgements

The first author wishes to announce that his stay at LAG was supported by Marie-Curie Fellowship program and the support for this research work has been provided by the GAČR project 102/05/H525: The Postgraduate Study Rationalization at Faculty of Electrical Engineering and Computer science VSB-TU Ostrava.

10 References

Application of fuzzy logic for gene expression analysis and classifications

Patryk Walawender, Łukasz Ćmielowski
Silesian University of Technology
cmielowski@tlen.pl, walek24@tlen.pl

Abstract: Thanks to microarray technologies we can measure genes expression and observe a lot of differences between tissues from various samples. This article examines the application of a fuzzy logic algorithm for DNA microarrays gene expressions. Our new approach will be presented and compared with similar approaches used in other publications. The algorithm proposed by us will be learning on training data sets and after it will use purchased knowledge to predict to which group unknown samples belong. Our results will be compared with results obtained by other algorithms spread in the Internet.

Keywords: gene expression, DNA microarrays, fuzzy logic, pattern discovery

1 Introduction

DNA microarray technology has been increasingly used in cancer research over the last few years. One of the major statistical tasks in studies involving these technologies is to find genes that are differentially expressed between two experimental conditions. The simplest example is to find genes that are up- or down-regulated in cancerous tissue relative to healthy tissue. Researchers have been trying for ages to improve cancer classification to determine which genes are differentially expressed between cancer and healthy tissue. Computer science and math help us to quickly find dependencies and make it possible to draw conclusions obtained by comparing ill and healthy samples with a huge amount of data.

In this paper we would like to present the application of fuzzy logic for gene expression analysis and classification. We are presenting an algorithm which consists of two parts - discovery and prediction. This method enables us to handle datasets where training samples exist and all the test samples belong to the known classes. By discovering we mean the study of both ill and healthy samples and finding differences. During this phase the algorithm will learn how to recognize if the tissue comes from ill or healthy sample group. This knowledge will be used in the second part of algorithm called prediction. We would like the algorithm to predict if a sample comes from an ill or healthy tissue leaning in the discovering phase.

The algorithm proposed by us is based on fuzzy logic rules selected in respect of efficiency and rate. The main problem was the investigation of the algorithms parameters to create a universal tool which could be used to study the majority of available samples. Our method and results are compared with similar publications [1, 2].
2 Methods

2.1 Discovering significant genes

The main aim of this part of algorithm is to find significant features of different classes using training set of samples with known classes.

Let $U$ be a set of tissue samples, let $G$ be a set of gene symbols, let $C$ be a set of class labels. Let $g(x)$ denote the value of expression of the gene $g$ in tissue sample $x$.

For two class labels $C_1$ and $C_2$ we have:

\[
U_1 \text{ – tissue samples from the class } C_1 \\
U_2 \text{ – tissue samples from the class } C_2
\]

where:

\[
U_1 \cup U_2 = U
\]

Every gene can be considered according to a qualitative levels such as up(u), neutral(n) and down(d) [3].

For each gene from each group $C_1$ and $C_2$ we define three corresponding functions $g_u, g_n$ and $g_d \ U \rightarrow [0,1]$ as (2) [2]:

Let ramp: $R^3 \rightarrow [0,1]$ be defined as:

\[
\text{ramp}(x, v, w) = \begin{cases}
0 & \text{if } x \leq v \\
1 & \text{if } x \geq w \\
\frac{x - v}{w - v} & \text{otherwise}
\end{cases}
\]

Associate with each gene $g$ a triplet $(g_{\text{min}}, g_{\text{median}}, g_{\text{max}}) \in R^3$

We define:

\[
\begin{align*}
    g_u &= \text{ramp}(g(x), g_{\text{median}}, g_{\text{max}}) \\
    g_d &= 1 - \text{ramp}(g(x), g_{\text{min}}, g_{\text{median}}) \\
    g_n &= 1 - \max(g_u(x), g_d(x))
\end{align*}
\]

We count three above values for all genes for both classes $C_1$ and $C_2$ separately, but $g_{\text{median}}, g_{\text{max}}$ and $g_{\text{min}}$ are counted always for one class. For example, for classes $C_1$ and $C_2$ we use during computation $g_{\text{median}}, g_{\text{max}}$ and $g_{\text{min}}$ values from the class $C_1$ or from class $C_2$ only.

For each gene $g$, $i = 1, \ldots, n$ - we compute:

\[
\begin{align*}
    \text{res}_1 &= \text{abs}(g_{u1}(i) - g_{u2}(i)) \\
    \text{res}_2 &= \text{abs}(g_{d1}(i) - g_{d2}(i)) \\
    \text{res}_3 &= \text{abs}(\text{res}_1 - \text{res}_2)
\end{align*}
\]

where 1 and 2 denote the class where the gene belongs.

After sorting results from $\text{res}_3$ we got the genes in order of their diversity. Thanks to it we can recognize the most significant genes which values differs considerably depending on class they belong.
2.2 Prediction origin of samples
After detecting the most significant genes they can be used to identify dependencies in unknown samples. In this part of algorithm test samples are being classified into known classes found in the point 2.1. Only one value describing how many significant genes should be considered is set.

for each \( g_k \), \( k = 1 : s \)
\( g_k \) - gene from set
\( s \) - denote number of significant genes

if the value of checked sample for gene \( g_k \) is higher than the maximum median of all classes value then the gene belongs into the same class as the counted median value

if the value of checked sample for gene \( g_k \) is lower than the minimum median of all classes value then the gene belongs into the same class as the counted median value

After all \( s \) iterations we compare how many times sample belonged to each group and then classify sample into a group with the highest value. Figure 1 shows the main idea of prediction algorithm for two classes.

![Fig. 1 The main idea of prediction algorithm](image)

3 Experiments
During our experiments we used two datasets of leukemia and lung cancer from[4]. The data have the following features:

**ALL-AML Leukemia**
Number of Instances: 38 training samples v.s. 34 testing samples
Number of Attributes: 7129 (all numeric)
Number of Classes: 2 (ALL v.s. AML)
Description: Training dataset consists of 38 bone marrow samples (27 ALL and 11 AML), over 7129 probes from 6817 human genes. Also 34 test samples data is provided, with 20 ALL and 14 AML.
**Lung Cancer**

Number of Instances: 32 training v.s. 149 testing  
Number of Attributes: 12533 (all numeric)  
Number of Classes: 2 (MPM v.s. ADCA)  
Description: Classification between malignant pleural mesothelioma (MPM) and adenocarcinoma (ADCA) of the lung. There are 181 tissue samples (31 MPM and 150 ADCA). The training set contains 32 of them, 16 MPM and 16 ADCA. The rest 149 samples are used for testing. Each sample is described by 12533 genes.

Above data sets are perfect for our attitude. We can teach our algorithm on the training samples and after that make an attempt of recognizing origin all samples from the testing groups.

**4 Results**

Figures 2 and 3 present the best 4 solutions for two datasets described in previous part 3. We can observe that found genes differ considerably depending on group they belong. The best 10 solutions for two groups are presented in Table 1. We have classified tests samples for ten characteristic genes and results are presented in Table 2.

![Gene Expression Graphs](image)

*Fig. 2* The best 4 solutions for ALL-AML Leukemia (27 ALL and 11 AML)
Table 1 The best 10 solutions for Leukemia and lung cancer data sets

<table>
<thead>
<tr>
<th>Leukemia</th>
<th>Lung Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>gene number</td>
<td>gene name</td>
</tr>
<tr>
<td>1</td>
<td>3320</td>
</tr>
<tr>
<td>2</td>
<td>2020</td>
</tr>
<tr>
<td>3</td>
<td>6218</td>
</tr>
<tr>
<td>4</td>
<td>1882</td>
</tr>
<tr>
<td>5</td>
<td>2288</td>
</tr>
<tr>
<td>6</td>
<td>4847</td>
</tr>
<tr>
<td>7</td>
<td>4499</td>
</tr>
<tr>
<td>8</td>
<td>2267</td>
</tr>
<tr>
<td>9</td>
<td>1926</td>
</tr>
<tr>
<td>10</td>
<td>5772</td>
</tr>
</tbody>
</table>

Fig. 3 The best 4 solutions for Lung Cancer (16 MPM and 16 ADCA)

Table 2 Results of classification for tests samples

<table>
<thead>
<tr>
<th></th>
<th>Leukemia</th>
<th>Lung Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of all tested samples</td>
<td>34</td>
<td>149</td>
</tr>
<tr>
<td>Number of positive classified samples</td>
<td>28</td>
<td>147</td>
</tr>
<tr>
<td>efficiency</td>
<td>82,4%</td>
<td>98,7%</td>
</tr>
</tbody>
</table>
5 Conclusion and Discussion

We have developed simple framework based on fuzzy logic for analyzing DNA genes expression data. Although using simple rules we achieved good results comparable with different works. Six solutions from ten detected by us for leukemia test samples belong to the best 50 solutions presented in [1]. Comparing to [2] we achieved better results for leukemia dataset – only 6 wrong predictions comparing to 7. Results received for lung data set are really satisfactory. Only two predictions from 149 samples are wrong which determine high result of effectiveness (98,7%). It seems that algorithm proposed by us ideally fits for analyzing gene expression. Figures 2 and 3 show that detected genes values differ considerably depending on group they belong to. Efficiency over 80% percent for leukemia test sample and almost 100% for lung cancer (Table 2) is even better than satisfactory. In the future we would like to improve efficiency of our algorithm probably by changing fuzzy rules and conduct more tests on available data sets.

6 References


Abstract: Currently 3D analysis has the great importance in current knowledge in image processing domain. 3D tests are made possible by the character of medical images from magnetic resonance MR or computer tomography CT. Spatial presentation of anatomical structures can be helpful in medical diagnostic and treatment of patients. Unfortunately available methods of 3D segmentation from serial images require manual interaction for more than one image (slice) which is uncomfortable and burdensome for the user. In the current study more automatic method is proposed. In this method semi-automatic analysis of the single image (slice) is sufficient. Proposed algorithm is an extension of the Live-Wire method.

1 Introduction

Image segmentation is very important problem in image processing particular in analysis of medical images. This operation allows to extract from image only regions of interest (in medicine: anatomical structures). Majority of the medical images from radiology examination are serial images (3D datasets, volumes). It makes possibility to segmentation of 3D solids and them visualization. In the current study extension of the Live-Wire method (also known as intelligent scissors) to 3D segmentation is proposed. Discussed algorithm depends on Live-Wire contour computing at the key slice (selected by the user) and then this information is automatically interpolated at remaining slices. Results are presented based on medical images from magnetic resonance of knee joint.

2 Related Work

Several extensions of the Live-Wire approach to 3D images (data volumes) have been addressed in the literature [2,3,4,5,7,8]. Unfortunately no one suggestion is automatic enough and accurate enough. One part of the segmentation methods [2,4,7] requires contours determination with user interaction at more than one slice (in the best case at two slices: the first where structure appears and the last where this structure disappears). It is burdensome and time consuming in exchange relatively high precision. Remaining methods [3,8] based on points interpolation. First, at selected slice, contour of segmented structure is computed with user interaction. Then a part of contour points (characteristic points i.e. inserted by the user) are interpolated at neighboring slices. Movement of selected points is automatic and depends on searching maximum value of image gradient into neighborhood of analyzed point. Next transferred points at each slice are connected based on 2D approach principles. This kind of methods is more automatic but with low precision. Because existing methods are failure searching of better approaches is justified and desirable. In the current study approach, which is compromise among discussed above kinds of methods, is presented.
3 2D Segmentation

Regardless of three-dimensional (3D) methodology two-dimensional (2D) segmentation has to be carry out. One of the 2D segmentation methods is Live-Wire approach from among edge-based segmentation methods. Live-Wire approach depends on: cost map calculation, image to graph transformation and optimal path searching. The most important element of the Live-Wire approach, which influence on accuracy, is cost map. It is a matrix with image size which describes edge properties at each image pixel. The low is cost matrix value, the more edge properties pixel has got. To find edge properties of the pixel different gradient features are applied. In traditional Live-Wire method [6] the first and the second derivative of image (and them modification like Laplacian zero-crossing) are utilized. But it is not one possible definition of the cost map. Perfect tool to image processing is wavelet transformation, thus wavelet definition of the cost map can be also efficient. Presented in [9] results prove this presumption. After cost map calculation graph is created from image and cost matrix values constitute weights of the graph vertices. When characteristic point (seed point) is interactively specified by the user spanning tree rooted at this point (vertex) is created and optimal path searching is started. Searching operation is the most time consuming element of the algorithm. It proceeds according to Dijkstra algorithm [1]. Each path in the graph is equivalent of any boundary in image. The number of founded paths is huge but only one is significant. Therefore another point specification – free point -- is necessary. Nevertheless all paths had to be earlier computed. To restrict number of computed paths and eliminate from calculation redundant paths Live-Wire on the fly [6] is applied. But to radically limit number of computed paths below method proposition, with cost map modification, can be utilized.

3.1 Cost Map Definition

To restrict number of searched paths Fuzzy C-Means clustering (FCM) can be helpful. In the case of grey level images (image with $2^8$ different grey level values, or with $2^{12}$ values in the case of medical images), after image clustering on C class transformed image with C different grey values is obtained. Easy to affirm that less than ten class (i.e. five classes, $c=5$) is sufficient to distinguish indispensable structures from image. Moreover it is possible to state that majority of these structures boundaries agree with boundaries between classes in classified image. It allows to consider only pixels situated on the boundaries. Therefore if only cost matrix elements situated near boundaries will be respect number of paths will be limited. Because cost matrix has to be complete (it is a matrix with image size) remaining elements not situated on boundaries should achieve large values. Then these points will not be taken into consideration in paths computing. There is also reasonable if the Live-Wire on the fly modification is simultaneous applied.

4 3D Extension

The goal of the presented method is contour moving of the object of interest from single slice interactively specified by the user (key slice) to each other slice from the set of serial images from the same radiology examination. Moving contour, which encloses segmented structure, is created according to Live-Wire method, therefore with user interaction, which depends on characteristic points (seed and free) specification. The traditional [6] or modified [9] Live-Wire method can be utilized to segmentation of this structure from key slice. This two-dimensional part of algorithm is necessary to 3D segmentation. Functionality of the proposed application to 3D segmentation is the following, Fig. 1.
The set of images from single MRI examination is demonstrated. User specifies single image (slice), named key slice, where segmented structure is the most clearly visible (or generally visible). Based on this slice 2D segmentation utilizing Live-Wire algorithm is realized. Next appointed contour is automatically (without user intervention) interpolated between neighboring slices. Contour moving is continued until lack of the structure at any slice will be achieved or until all slices are analyzed. Crucial to method correctness is contour interpolation.

4.1 Contour Interpolation

In order to move contour from actual to neighboring slice individual contour points have to be interpolated. Selection of the points can be distinct e.g.: there are points where the first derivative of gradient direction is maximal. In the discussed case all contour points are taken into consideration because continuation of the method is not numerically complex and it allows to obtain more accurate results. Method of contour interpolation utilized fuzzy C-means algorithm. In the first step key slice, after contour determination with Live-Wire algorithm, is classified. The number of class is dynamically determined. It starts from two classes and pass until boundary of classified image will be covered with Live-Wire contour. All classes situated inside Live-Wire contour are joined to create binary image. This image has high values (white pixels) in the place of segmented structure, surrounded with Live-Wire contour. Next centroid of this area – center of geometry – is computed. This specific point is moved at next and previous slices – on condition that they exist – or only at one of them (so that operation of points interpolation proceeds in the one direction).

Because proposed algorithm, during searching of segmented structure, takes into consideration only pixels situated on boundaries between areas from classified version of image, each slice from set of serial images is classified at the same, dynamically determined, number of classes. Centroid, which is moved at neighboring slice or slices, becomes central point for it or them. From central point (points) radial profiles are computed. Along each profile – starting from central point – boundary of different class is searched. The first, encountered point of other class is treated as a new contour point, situated on neighboring – relative to key – slice. This operation of point moving is realized from key slice and is propagated into two directions: increased and decreased number of slices. Actually searched contour becomes basic contour (or key contour) for its neighboring slice in suitable direction.
Algorithm stops when either end of set of slices is reached or at any slice lack of segmented structure is occurred. The most difficult task in this algorithm is estimation of number of classes utilized to image segmentation. In the current study criterion of majority was applied (majority of contour points have to be covered with points of boundary into classified image). It is not optimal but in the case of expressive structures it is sufficient. Generalization of the method for other types of objects is planed.

Fig. 1. 3D segmentation – the segmented structure is seen at each slice (darker contour in the left lower corner was made interactively by the user)
5 Results

Proposed method was tested based on medical images from magnetic resonance (MRI) of knee joint. There are two possibilities: segmented structure is seen at each slice from full set of images or only at some of them. These two cases was presented adequately in Fig. 2 and Fig. 3.
Because in Fig. 2 segmented structure is seen at each slice algorithm finds contours at each of them. But in Fig. 3 desirable anatomical structure is not seen at last slide, thus no one contour at this slide has been found. Presented approach possess limitations and disadvantages. Segmented objects have to characterize compact structure. This structure should be enclosed by single, closed contour and have to include in consecutive slices (number of slices does not matter). Undisputed advantage of the method is its automation. Thanks to it only user interaction at key slice is required.

6 References

Comparative Analysis of Classical PI Speed Controller and Fuzzy Controller Used in Direct Torque Control of Induction Motor

Lamia YOUB, Aurelian CRACIUNESCU

POLITEHNICA University of Bucharest, Electrical Engineering Faculty, 313, Splaiul Independentei, 060042 Bucharest, Romania, youblamia@yahoo.fr

Abstract- One of the frequently discussed applications of artificial intelligence in motion control is the replacement of a classical PI speed controller with a fuzzy logic (FL) speed controller. In this paper, a complete and rigorous comparison is made between the two types of controllers. The fuzzy controller can be used not only as a substitute but as a complimentary controller to the classical PI controller. DTC was simulated for various cases. Results obtained show that the fuzzy logic technique can provide better speed control performance of DTC.

1 Introduction

Traditional direct torque control system of induction motor [1] introduces conventional PI regulator in speed loop, it is proved that the low precision of the speed regulator debases the performance of the whole system. This paper establishes a fuzzy speed PI regulator [2], which applies the principles and method of fuzzy logic to adjust the proportional coefficient $k_p$ and integral coefficient $k_i$ of the PI regulator on-line. The results of simulation show that the fuzzy speed regulator can ensure swift speed response, small overshooting, and high steady speed precision both in high and low speed. Additionally, the proposed regulator improves the better speed control performance of DTC. In traditional direct torque control system, speed regulator is widely based on conventional PI regulator, which has to depend on precise math model of the subject. Since direct torque control technology does not decouple the induction motor model, it is difficult to confirm the relationship between set speed and torque, when induction motor variables have changed or environment disturbances have occurred. Furthermore, the conventional parameter-fixed speed regulator cannot easily achieve swift response, small overshooting and fine speed control precision in a wide speed.

2. DTC Strategies

This method is based on maintaining constant amplitude and phase of the stator currents [3], avoiding electromagnetic transients. It is possible to control directly the stator flux and torque by selecting the appropriate inverter state [1], [5]. An induction machine can be modeled with stator current and flux in reference $(\alpha, \beta)$ as state variable by the following equations.

$$\dot{X} = AX + BV_s$$

(1)

Where:
The induction stator flux and torque are given by:

\[
\phi_s = \int (v_s - R_s i_s) dt \\
T_e = p_i (\phi_{s_{\alpha}} i_{s_{\beta}} - \phi_{s_{\beta}} i_{s_{\alpha}})
\]

The estimated values of the stator flux and torque are compared to their prescribed values \(\Phi_{s\text{ref}}, T_{\text{ref}}\) respectively. Switching states are selected according to switching table selector, where \(S_T\) is the stator flux modulus after the hysteresis block and \(S_\phi\) is the torque error after the hysteresis block.

<table>
<thead>
<tr>
<th>Flux</th>
<th>Torque</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>(S_T=1)</td>
<td>(S_\phi=1)</td>
<td>(V_2)</td>
<td>(V_3)</td>
<td>(V_4)</td>
<td>(V_5)</td>
<td>(V_6)</td>
<td>(V_1)</td>
</tr>
<tr>
<td>(\Phi_T=0)</td>
<td>(V_7)</td>
<td>(V_0)</td>
<td>(V_7)</td>
<td>(V_0)</td>
<td>(V_7)</td>
<td>(V_0)</td>
<td>(V_0)</td>
</tr>
<tr>
<td>(S_\phi=-1)</td>
<td>(V_6)</td>
<td>(V_1)</td>
<td>(V_2)</td>
<td>(V_3)</td>
<td>(V_4)</td>
<td>(V_5)</td>
<td>(V_0)</td>
</tr>
<tr>
<td>(S_T=1)</td>
<td>(S_\phi=1)</td>
<td>(V_3)</td>
<td>(V_4)</td>
<td>(V_5)</td>
<td>(V_6)</td>
<td>(V_1)</td>
<td>(V_2)</td>
</tr>
<tr>
<td>(\Phi_T=0)</td>
<td>(V_0)</td>
<td>(V_7)</td>
<td>(V_0)</td>
<td>(V_7)</td>
<td>(V_0)</td>
<td>(V_7)</td>
<td>(V_7)</td>
</tr>
<tr>
<td>(S_\phi=-1)</td>
<td>(V_5)</td>
<td>(V_6)</td>
<td>(V_1)</td>
<td>(V_2)</td>
<td>(V_3)</td>
<td>(V_4)</td>
<td>(V_4)</td>
</tr>
</tbody>
</table>

In the basic DTC scheme, as indicated in Fig.1, the stator flux linkage is estimated by means of integration of the terminal voltage minus the ohmic voltage drop on the stator resistance, as described in (5) [3]. The error between the estimated torque \(T\) and the reference torque \(T^*\) is the input of the three level hysteresis comparator where the error between estimated stator flux magnitude \(\Phi_s\) and the reference stator flux magnitude \(\Phi_{s^*}\) is the input of a two level hysteresis comparator. The block diagram of direct torque control system for an induction motor is shown in Fig. 1. In this system, the control reference frame is stationary (fixed to the stator) and space vector notation is used to represent the variables. The motor torque and stator flux amplitudes are controlled by two independent hysteresis controllers. The feedback signal, \(T_e\) and \(\Phi_e\), are computed from stator voltages and currents. The stator flux space vector \(\Phi_s\) is obtained by integrating the motor space vector [4]:

\[
X = \begin{bmatrix} i_s \\ \phi_s \end{bmatrix}, \quad X = \begin{bmatrix} i_i \\ \phi_i \end{bmatrix}
\]

\[
A = \begin{bmatrix} -\frac{R_s}{L_s} + \frac{1-\sigma}{\sigma T_r} M & \frac{M - \frac{1}{T_r}}{M} \\ \frac{M}{T_r} & -1 + j\omega \end{bmatrix}
\]

\[
B = \begin{bmatrix} \frac{1}{L_s L_r} \sigma & \frac{1}{L_r} \\ 0 & 0 \end{bmatrix}
\]

\[
j = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad T_s = \frac{L_s}{R_s}, \quad T_r = \frac{L_r}{R_r}
\]
2.1 Design of the fuzzy speed PI regulator

In the DTC system, the discrete expression of the conventional speed PI regulator is showed as Equation (8):

\[
T_d^*(k) = K_p e_o(k) + K_i \sum_{j=1}^{k} e_o(j)
\]

(8)

Where, \( T \) is sampling period \( e_o(k) \), is speed error signal, \( T_d^*(k) \) is set value of torque output, \( K_p \) is proportional coefficient, and \( K_i \) is integral coefficient. Equation (8) infers that control effects can easily be acquired at different speed demand, by dynamically adjusting \( K_p \) and \( K_i \) according to the input speed variable.

2.1.1 Selection of fuzzy subsets

The fuzzy controller is a system with 2-inputs, 1-output, and the inputs for fuzzy controller are the error and the integral of the response, the output of the fuzzy controller is the response of the system. The results of this type of fuzzy controller have shown that the performance of the fuzzy PI controller is comparable to the performance of the classical PI-controller. The structure of the previous type of controller contained first of all a static fuzzy section, the output of which was usually (with the PI type). In the presented structure there is first a dynamic section in the form of a continuous PI member, and then the static non-linear fuzzy section of the controller. The inputs includes 7 fuzzy subsets {negative positive, negative medium, negative small, zero, positive small, positive medium,...}, its membership function is shown as fig.2. Their membership functions are shown as Fig. 3 and Fig. 4, respectively.

Fig. 1 Basic direct torque control scheme for ac motor drives (DTC)

Fig. 2 Inputs membership functions
3.1 Establishment of fuzzy control rule

The performance of a fuzzy controller is defined by the dependence between the controller inputs and outputs, this dependence being described not by analytical equations, but by rules of the following type: IF $e_1$ is. AND $e_2$ is. AND $e_3$ is. THEN $u$ (or $\Delta u$) is (1). In these rules, close to human language description, $e_i$ is ith input variable and $u$ is the output of the fuzzy controller static part (or $\Delta u$ is the gain of the output variable). As the operation of a fuzzy controller is based on qualitative knowledge about the system being controlled, it is first important to identify these properties. Our selection of the fuzzy controller input variables is based on Fig. 1 which represents the response of the controlled system output $y$ to the jump of the desired value of control variable $r$.

<table>
<thead>
<tr>
<th>$\Delta e$</th>
<th>NB</th>
<th>NS</th>
<th>NM</th>
<th>ZO</th>
<th>PS</th>
<th>PM</th>
<th>PB</th>
</tr>
</thead>
<tbody>
<tr>
<td>e Out</td>
<td>NB</td>
<td>NB</td>
<td>NM</td>
<td>NM</td>
<td>NS</td>
<td>NS</td>
<td>ZO</td>
</tr>
<tr>
<td></td>
<td>NS</td>
<td>NB</td>
<td>NM</td>
<td>NM</td>
<td>NS</td>
<td>ZO</td>
<td>PS</td>
</tr>
<tr>
<td></td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NS</td>
<td>NS</td>
<td>ZO</td>
<td>PS</td>
</tr>
<tr>
<td></td>
<td>ZO</td>
<td>NM</td>
<td>NS</td>
<td>ZO</td>
<td>PS</td>
<td>PS</td>
<td>PM</td>
</tr>
<tr>
<td></td>
<td>PS</td>
<td>NS</td>
<td>ZO</td>
<td>PS</td>
<td>PS</td>
<td>PM</td>
<td>PM</td>
</tr>
<tr>
<td></td>
<td>PM</td>
<td>NS</td>
<td>ZO</td>
<td>PS</td>
<td>PM</td>
<td>PM</td>
<td>PB</td>
</tr>
<tr>
<td></td>
<td>PB</td>
<td>ZO</td>
<td>PS</td>
<td>PS</td>
<td>PM</td>
<td>PM</td>
<td>PB</td>
</tr>
</tbody>
</table>

The following plots are showings an example of the member functions using for the fuzzy controller to proof the performance of the fuzzy controller, the results of the fuzzy controller system using different membership functions are compared with those of the classical PI controller and fuzzy controller as a compliment to the classical PI controller.
Table 3 Response of input with variable functions for PI fuzzy and PI-PI fuzzy

<table>
<thead>
<tr>
<th>Functions</th>
<th>T1</th>
<th>T2</th>
<th>Overshoot</th>
</tr>
</thead>
<tbody>
<tr>
<td>PID</td>
<td>0.4</td>
<td>0.8</td>
<td>4.73</td>
</tr>
<tr>
<td>TRIANG</td>
<td>0.424</td>
<td>0.422</td>
<td>0.29</td>
</tr>
<tr>
<td>TRAPEZ</td>
<td>0.412</td>
<td>0.354</td>
<td>0.18</td>
</tr>
<tr>
<td>DSIGMF</td>
<td>0.373</td>
<td>0.428</td>
<td>0.26</td>
</tr>
<tr>
<td>ZMF</td>
<td>0.3407</td>
<td>0.422</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Table 4 Response of input and output with variable functions for PI fuzzy and PI-PI fuzzy

<table>
<thead>
<tr>
<th>Functions</th>
<th>T1</th>
<th>T2</th>
<th>Overshoot</th>
</tr>
</thead>
<tbody>
<tr>
<td>PID</td>
<td>0.4</td>
<td>0.8</td>
<td>4.73</td>
</tr>
<tr>
<td>TRIANG</td>
<td>0.424</td>
<td>0.422</td>
<td>0.29</td>
</tr>
<tr>
<td>TRAPEZ</td>
<td>0.354</td>
<td>0.412</td>
<td>0.18</td>
</tr>
<tr>
<td>DSIGMF</td>
<td>0.428</td>
<td>0.373</td>
<td>0.26</td>
</tr>
<tr>
<td>ZMF</td>
<td>0.3407</td>
<td>0.422</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Table 5 Response for variable functions for PI- fuzzy and PI-PI fuzzy

<table>
<thead>
<tr>
<th>Functions</th>
<th>T1</th>
<th>T2</th>
<th>Overshoot</th>
</tr>
</thead>
<tbody>
<tr>
<td>PID</td>
<td>0.4</td>
<td>0.8</td>
<td>4.73</td>
</tr>
<tr>
<td>TRIANG</td>
<td>0.422</td>
<td>0.424</td>
<td>0.29</td>
</tr>
<tr>
<td>TRAPEZ</td>
<td>0.448</td>
<td>0.354</td>
<td>0.18</td>
</tr>
<tr>
<td>DSIGMF</td>
<td>0.433</td>
<td>0.373</td>
<td>0.29</td>
</tr>
<tr>
<td>ZMF</td>
<td>0.3407</td>
<td>0.422</td>
<td>0.29</td>
</tr>
</tbody>
</table>
4. Simulation results

In this paper, we apply Simulink tool in Matlab software to make simulation for speed control system. Motor’s parameters for simulation are: $P_n=4\,\text{kW}$, $U_n=220/380\,\text{V}$, $R_s=10\,\Omega$, $R_m=6.3\,\Omega$, $L_s=0.6550\,\text{H}$, $L_r=0.6520\,\text{H}$, $L_m=0.612\,\text{H}$, $J=0.03\,\text{kg.m}^2$, $P=2$, the simulation results are shown as Fig.6, 7, 8 response speed in fuzzy control system and robust fuzzy controller the fig. (6, 7, 8) compares the performance of the stable fuzzy controller and a compliment to the fuzzy controller. In all cases the fuzzy control designs get the job done but with different performance characteristics. The robust fuzzy controller is the most of the robustness.

5. Conclusion

The paper, based on the DTC of induction motor, proposes using the fuzzy controller speed regulator as the speed regulator in the DTC system. The introduction of the PI fuzzy speed controller improves the speed adjustment capability of the DTC system. It ensures speed response small overshooting and fine precision of speed. Fuzzy speed regulator proposed in the paper not only improves the whole control performance of DTC system, but also enhances the robustness of the system with fuzzy controller.

6. References

Localization of the Posterior Cruciate Ligament on the T1-weighted MR Knee Images

Piotr Zarychta, Anna Zarychta-Bargiela
Silesian University of Technology
piotr.zarychta@polsl.pl

Abstract: I would like to present a new method of the posterior cruciate ligament localization on the T1-weighted MR knee images. This localization allows to reduce number of slices used in 3D rendering. Each slice (T1-weighted MR knee images Fig. 1) is subjected to two-dimensional entropy and energy measurements of fuzziness. Then, a fuzzy c-means clustering technique modified by introducing additionally to the objective function a median estimator is implemented. The class reflecting soft tissue is subjected to similarity measures which compare a local image intensity. A detailed local analysis, based on image profile models, yields a 3D region of interest (ROI) which includes the posterior and anterior cruciate ligament. Finally, fuzzy techniques applied to the ROI histogram analysis enhances the contrast and views the 3D region.

1 Introduction

Posterior and anterior cruciate ligament (PCL/ACL) are anatomical structures, which are (together with collateral ligament) responsible for the knee stability. Cruciate ligaments are main stabilizers of the knee joint in the sagittal plane; they also contribute in stabilisation in coronal and transverse planes. Together with the shape of articular surface, muscles and contact forces ensure proper arthrokinematics. During passive motion of the knee cruciates help to change rolling into sliding movements and during active motion resist translations and reduce shear forces. Cruciates control rotational movements in the flexed knee and together with collateral ligaments ensure rotational stability of the extended knee. For that reason cruciate ligaments belong to the group of anatomical structures, which are frequently susceptible to injury, especially in the case athletes. Much has been written regarding general treatment and various surgical procedures of the destabilized knee since only a few musculoskeletal conditions have stimulated as much controversy as an injury to the ACL and PCL. Once considered the beginning of the end of the normal knee function nowadays appears...
to be treatable. The current prognosis with appropriate treatment appears improved at least over
the short term. However the problem should not be considered solved as long term outcome
studies proof that degenerative joint disease can be delayed or prevented. The success of
ligaments reconstructive procedure depends on many factors mainly accurate diagnosis based
on the localization and visualization of the anatomical structures of the cruciate ligament.

2 The entropy measure of fuzziness

In order to introduce measures of fuzziness in image quantitative description, let us first
discuss a concept of a fuzzy image. This concept is based on the idea of a fuzzy signal,
introduced by Czogala and Leski in [1, 2]. Let us consider an image X(N,M) at the size of
NxM with pixels I(n,m) and n=1,2,…,N m=1,2,…,M. Moreover, the image is scanned within
a window at the size of (2k+1)x(2k+1). The concept of a fuzzy image derived from the
original image is based on two assumptions: if no fuzzy uncertainty is considered in the
image, a fuzzy image pixel I(n,m,k) is reduced to a real number I(n,m) referred to as a
singleton. The measure of fuzziness is equal to zero and the information is maximal; the
measure of fuzziness increases if the original image changes, as higher the dynamic, as
smaller amount of information is conveyed by the image. In order to construct a fuzzy image
from crisp image, the pixels within the window are sorted into an increasing order is shown in
the Fig. 2. The median value is located at the (k+1,k+1) position within the window, ie.
I_{MED}(n,m,k)=I_{k+1,k+1}(n,m,k). On the basis of the set of pixels put in the increasing order and
the median, the membership function μ is defined. First we assume, that
\[ \mu_{n,m,k}(I_{\text{min}}(n,m,k))=0, \mu_{n,m,k}(I_{\text{max}}(n,m,k))=0 \text{ and } \mu_{n,m,k}(I_{\text{MED}}(n,m,k))=1. \]
The remaining elements are defined according to the following formula
\[ \mu_{n,m,k}(I_{i,j}(n,m,k)) = \frac{(2k-d_{i,j}+1)/(2k+1)}, \]
where \( d_{i,j}=|n-i|+|m-j| \). In order to discriminate against a certain level of membership \( \lambda \in [0,1] \),
the Heaviside pseudofunction
\[ I(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \]
is introduced. Implementation the Heaviside function is defined as
\[ \mu_{n,m,k}^\lambda(I_{i,j}(n,m,k)) = \mu_{n,m,k}(I_{i,j}(n,m,k))1(\mu_{n,m,k}(I_{i,j}(n,m,k)) - \lambda). \]
The introduced of the \( \lambda \) level into the measurement of fuzziness allows the insignificant
membership degree to be removed. A concept of the entropy measure of fuzziness has been
introduced in [3] and implemented in fuzzy signal concept in [1, 4, 5]. The entropy measure
of fuzziness is a mapping from the set of all fuzzy subsets of a base set X into the nonnegative

![Sorted pixels within the window](image-url)
reals. It can be expressed as
\[
H(A, \lambda) = \int_{X} \frac{\lambda}{\lambda + \mu(A(x))} dv,
\]
where \(A: X \rightarrow [0,1]\) is any \(\nu\)-measurable function, \(dv = dx\) or \(dv = p(x)dx\), where \(p(x)\) stands for a probability density function; \(h:[0,1] \rightarrow \mathbb{R}_+\) is an increasing function in \([0, 0.5]\), a decreasing function in \([0.5, 1]\), and \(h(0) = h(1) = 0\); \(F: \mathbb{R}_+ \rightarrow \mathbb{R}_+\) is an increasing function and \(F(z) = 0\) if \(z = 0\). Substituting (3) into (4) we obtain
\[
H\left(\mu_{n,m,k}^\lambda I_{i,j}\right) = F_{1} \left( \sum_{i=1}^{2k} \sum_{j=1}^{2k} h_{\lambda} \left( \mu_{n,m,k}^\lambda I_{i,j}\right) \right) \cdot p(I_{i,j}) \cdot \Delta I_{i,j},
\]
where \(I_{i,j}\) denote \(I_{i,j}(n,m,k)\), and \(h_{\lambda}(z)\) is defined as
\[
h_{\lambda}(z) = \begin{cases} h(z) & \text{if } z \in (\lambda, 1-\lambda) \\ 0 & \text{otherwise} \end{cases}
\]
and \(\Delta I_{i,j}(n,m,k)\) is a gradient between neighbour pixel values as marked in the Fig. 2, and the probability density function \(p(I_{i,j}(n,m,k))\) is obtained from a histogram, by dividing the number of pixels by \(2k+1\).

### 3 Median filtering

Standard Fuzzy C-Means (FCM) algorithm [6] is widely used in many clustering approaches. Its advantages include a conceptual and computational simplicity and the ability to model uncertainty within the data. FCM has also several weaknesses. It does not incorporate spatial context information which makes it sensitive to noise and images artifacts. In this paper is used a modified method of FCM to the clustering. The FCM objective function is modified by adding a second term that formulates a spatial constraint based on the median estimator. Median of a set \(\{x_1, ..., x_n\}\) is an M-estimator of location, with a cost function [7] given as
\[
M(\zeta) = \sum_{i=1}^{N} |x_i - \zeta|.
\]
Median of an ordered data set \(A = \{x_1, ..., x_N\}\) is defined as follows
\[
\text{Median}(A) = \left\{ \begin{array}{ccc} x_{(n+1)/2} & & n = 1, 3, 5, ... \vspace{0.2cm} \\ 0.5 \cdot (x_{(n+0.5)} + x_{n+0.5}) & & n = 2, 4, 6, ... \end{array} \right.
\]
In image processing approaches an implementation of median running-window filtering, replaces each data sample by its spatial neighborhood function. Neighborhood function is defined
\[
\text{MEDF}(x, Z) = \text{median}(S),
\]
where \(S = \text{neighborhood}(x, Z)\) and \(Z\) determines the size of the mask. Standard FCM is defined as (11) and is derived to minimize the objective function with respect to the membership function \(u_n\) and the center \(v_i\) for a given fuzzyfication level \(m\) (where \(1 \leq m < \infty\)).
\[
M(U, V) = \sum_{i=1}^{N} \sum_{n=1}^{N} u_{n}^{m} \|x_{n} - v_{i}\|^{2},
\]
where \(x_n = \{x_1, ..., x_k\}\) and \(x_n, v_i \in F^k\). \(v_i\) is called a prototype and \(F^k\) is a feature space. Chen and Zhang [8] have modified the objective function of FCM by introducing an element, that depends on the mean value of neighbouring pixels. Since the modification propagates features of a mean filtered image into the clustering results, blurred edges is one of the most
important disadvantages of the method. In order to reduce the drawback, median estimator has been added into the objective function [9, 10]

\[
M(U, V) = \sum_{i=1}^{N} \sum_{n=1}^{N} u_{nm}^{m} \left\| x_{n} - v_{i} \right\|^{2} + a \| MEDF(x_{n}, Z) - v_{i} \|^{2} \].
\] (12)

In the Fig.3 is shown the original signal from data base clinical hospital MR knee T1-weighted with entropy measures of fuzziness (Fig. 3a) and entropy measures of fuzziness with the FCM median modified (Fig. 3b).

### 4 Membership function

On the basis of analysis each profile in all slices of the MR knee T1-weighted after fuzzification and fuzzy clustering (Fig. 3b), is determined a main axis running along thighbone and tibia. The main axis is determined according to following formula

\[
\text{th} = \min \left[ L_{pr}(k) \right], \quad \text{for all} \quad L_{pr}(k) \neq 0,
\] (13)

where \(L_{pr}(k)\) denote number of stripe (nonzero value of the pixel intensity) in the profile of \(k\). In case of several profiles, which meet the condition (13), profile with the lowest column number is chosen (Fig. 4a). The next step in the process of the determining membership function requires executing a centering and superposition operation (Fig. 4b). Profiles analysis
after the centering and superposition operation permits creating the membership function. This stage gives information, which are indispensable for the formation of the membership function (location and size of the stripes in the main axis). On the basis of this analysis membership function has been determined (Fig. 5)

\[ \mu(i) = \sum_{n=1}^{3} f_{G}(i, \sigma_n, k_n), \quad (14) \]

where \( f_{G} \) denotes a Gaussian function. The Gaussian function depends on two parameters \( \sigma \) (standard deviation in the specific range) and \( k \) (average number of column in the specific range) as given by

\[ f_{G}(i, \sigma_n, k_n) = \exp \left( \frac{-(i-k_n)^2}{2\sigma_n^2} \right). \quad (15) \]

In my experiment standard deviation and average number of column were equal \( \sigma_1 = 7 \), \( k_1 = 25 \), \( \sigma_2 = 19 \), \( k_2 = 132 \), \( \sigma_3 = 5 \), \( k_3 = 222 \) respectively.

### 5 Region of interest (ROI)

In purpose to determine edges of the ROI including the posterior cruciate ligament, the following rules have been formulated:

- **Left edge** (axis I in the Fig. 6b)

\[ \min_k \left[ \min_j \bar{\mu}(i), pr_j(k) \right], \quad (16) \]

where \( k \) denotes a profile number in the slice, \( j \) is a slice number in the image, \( \bar{\mu} \) denotes a complement of the membership function (Fig. 5b), \( pr_j(k) \) is a \( i \)-value of a \( k \)-profile;

- **Upper edge** (axis II in the Fig. 6b)

\[ \max_u \left[ d_u \right], \quad (17) \]

where \( d_u \) denotes a distance between I axis (Fig. 5b) and right edge of thighbone for \( u \)-profile;

- **Lower edge** (axis III in the Fig. 6b)

\[ \max_p \left[ d_p \right], \quad (18) \]

where \( d_p \) denotes a distance between I axis (Fig. 5b) and right edge of tibia for \( p \)-profile.

**Right edge** (axis IV in the Fig. 6b) has been determined on the basis of the axis I shift by distance \( \max_p \left[ d_p \right] \). The determining profiles permit on constructing a region of interest (ROI), which includes posterior cruciate ligament (Fig. 5c).
Fig. 5. Localization of the PCL in the knee, (a) slice no.12 (Fig.3b), (b) marked axis, (c) ROI

6 Numerical Results

The methodology has been tested on 58 clinical T1-weighted MR knee studies. In 84% cases the region of interest (ROI) including the posterior cruciate ligaments has been extracted correctly. The remain regions, mostly with a severe ligament injury, require a manual enlargement. Introduced application software allows for an automatic extraction of a the ROI extracted from T1- and T2-weighted MR knee studies. Localization of the PCL on the T1- and T2-weighted MR knee images allows also the three-dimensional visualization in the process of the computer aided diagnosis of the cruciate ligament. In my opinion the proposed method of the posterior cruciate ligament localization in T1-weighted MR knee images seems to be very effective and promising.

7 References

Registration of the T1- and T2-weighted MR Knee Images

Piotr Zarychta, Anna Zarychta-Bargiela
Silesian University of Technology
piotr.zarychta@polsl.pl

Abstract: In this paper is shown a new method of the automatic registration of T1- and T2-weighted MR knee images. This method is based on the entropy and energy measures of fuzziness and can be used in localization process of cruciate ligament. First, two sequences (T1- and T2-weighted) are converted to a fuzzy representation. Then, the entropy and energy measures are employed in the NCC (normalized cross correlation) and GD (gradient difference) methods. The alignment based on energy and entropy fuzzy measures shows a significant improvement in comparison with the implementation of the original image.

1 Introduction

Image registration is a difficult and complex problem. According to the database of the Institute of Scientific Information, in the last 15 years more than 1000 papers were published on the topic of image registration [1]. For example clinical diagnosis and treatment verification is often supported by imaging modalities. These modalities provide functional (SPECT-Single Photon Emission Computed Tomography, PET-Positron Emission Tomography, MRS-Magnetic Resonance Spectroscopy) and anatomical (MRI-Magnetic Resonance Imaging, US-Ultrasound, X-ray imaging with CT-computed tomography) information. The functional images provide information of the metabolic processes of the human body. The anatomical images provide information of the anatomic structure of the human body. In many instances it is necessary to integrate the information obtained from two or more studies of the same patient. But the differences in patient positioning, different image acquisition parameters require the registration of these images before overlapping them. Procedure of the registration is usually required in following fields: medical image analysis (detection and localization of anatomical structures of the human body), remotely sensed data processing (geography, cartography, geology, oceanography, agriculture, oil and mineral exploration, monitoring with localization, detection and comparison, weather forecasting), computer vision and pattern recognition (segmentation, object recognition, shape reconstruction) [2, 3]. In medical application, registration process must be automatic, reliable registration of T1- and T2-weighted MR knee images (fig. 1). This procedure is used in locali-
zation and 3-dimensional visualization of the cruciate ligament. The image analysis in knee ligament injuries is often performed on the T1- and T2-weighted MR images simultaneously. Slices of the both series are consulted at the same time. Thus, the image display should view corresponding images of the both series. A registration procedure is required in order to view matching slices next to each other. Two functions have been developed. First, a concept of a fuzzy image is introduced. Second, entropy and energy measures of fuzziness are employed as image matching indicators.

2 A fuzzy image concept

In order to introduce measures of fuzziness in image quantitative description, let us first discuss a concept of a fuzzy image. This concept is based on the idea of a fuzzy signal, introduced by Czogala and Leski in [5, 6]. Let us consider an image X(N,M) at the size of NxM with pixels I(n,m) and n=1,2,...,N m=1,2,...,M. Moreover, the image is scanned within a window at the size of (2k+1)x(2k+1). The concept of a fuzzy image derived from the original image is based on two assumptions: if no fuzzy uncertainty is considered in the image, a fuzzy image pixel I(n,m,k) is reduced to a real number I(n,m) referred to as a singleton. The measure of fuzziness is equal to zero and the information is maximal; the measure of fuzziness increases if the original image changes, as higher the dynamic, as smaller amount of information is conveyed by the image. In order to construct a fuzzy image from crisp image, the pixels within the window are sorted into an increasing order is shown in the fig. 2. The median value is located at the (k+1,k+1) position within the window, ie. \( I_{\text{med}}(n,m,k)=I_{k+1,k+1}(n,m,k) \). On the basis of the set of pixels put in the increasing order and the median, the membership function \( \mu \) is defined. First we assume, that \( \mu_{n,m,k}(I_{\text{min}}(n,m,k))=0 \), \( \mu_{n,m,k}(I_{\text{max}}(n,m,k))=0 \) and \( \mu_{n,m,k}(I_{\text{med}}(n,m,k))=1 \). The remaining elements are defined according to the following formula

\[
\mu_{n,m,k}(I_{i,j}(n,m,k)) = \frac{(2k-d_{i,j}+1)/(2k+1),
\]

where \( d_{i,j} = |n-i|+|m-j| \). In order to discriminate against a certain level of membership \( \lambda \in [0,1] \), the Heaviside pseudofunction

\[
1(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}
\]

is introduced. Implementation the Heaviside function is defined as

\[
\mu_{n,m,k}^\lambda(I_{i,j}(n,m,k)) = \mu_{n,m,k}(I_{i,j}(n,m,k))1(\mu_{n,m,k}(I_{i,j}(n,m,k)) - \lambda).
\]

The introduced of the \( \lambda \) level into the measurement of fuzziness allows the insignificant membership degree to be removed.

3 The entropy and energy measures of fuzziness

A concept of the entropy measure of fuzziness has been introduced in [7] and implemented in fuzzy signal concept in [5, 8, 9]. The entropy measure of fuzziness is a mapping from the set of all fuzzy subsets of a base set X into the nonnegative reals. It can be expressed as

\[
H(A, \lambda) = \int_X h_\lambda^A(A(x))dv,
\]

where A:X→[0,1] is any v-measurable function, dv=dx or dv=p(x)dx, where p(x) stands for a probability density function; h:[0,1]→R_+ is an increasing function in [0, 0.5], a decreasing function in [0.5, 1], and h(0)=h(1)=0; F:R_+→R_+ is an increasing function and F(z)=0 if z=0.
Substituting (3) into (4) we obtain

\[ H\left(\mu_{n,m,k}^i, I_{j,i}\right) = F_1 \left( \sum_{i=1}^{2k} \sum_{j=1}^{2k} h_\lambda \left( \mu_{n,m,k}^i, I_{j,i} \right) \cdot p\left( I_{j,i} \right) \cdot \Delta I_{j,i} \right), \] (5)

where \( I_{j,i} \) denote \( I_{j,i}(n,m,k) \), and \( h_\lambda(z) \) is defined as

\[ h_\lambda(z) = \begin{cases} h(z) & \text{if } z \in (\lambda, 1-\lambda) \\ 0 & \text{otherwise} \end{cases}, \] (6)

and \( \Delta I_{j,i}(n,m,k) \) is a gradient between neighbour pixel values as marked in the fig. 2, and the probability density function \( p(I_{j,i}(n,m,k)) \) is obtained from a histogram, by dividing the number of pixels by \( 2k+1 \). The probability density function, as defined above, serves also as a basis for the energy extraction. The energy measure is expressed as

\[ E(A, \lambda) = \int_X f_\lambda(A(x)) dv, \] (7)

where

\[ f_\lambda(z) = \begin{cases} f(z) & \text{if } z \in (\lambda, 1) \\ 0 & \text{otherwise}. \end{cases} \] (8)

The final formula for the energy measure can be written as

\[ E(\mu_{n,m,k}^i, I_{j,i}) = F_2 \left( \sum_{i=1}^{2k} \sum_{j=1}^{2k} f_\lambda \left( \mu_{n,m,k}^i, I_{j,i} \right) \cdot p\left( I_{j,i} \right) \cdot \Delta I_{j,i} \right), \] (9)

where \( I_{j,i} \) denote \( I_{j,i}(n,m,k) \), \( \Delta I_{j,i}(n,m,k) \) and \( p(I_{j,i}(n,m,k)) \) are in (5).

Fig. 2 Sorted pixels within the window

### 4 Similarity measures

Next stage in computer aided diagnosis (CAD) is a similarity evaluation of the current images. Here are two main groups of similarity measures: feature-based and intensity-based. Feature-based measures make some processing with the images first in order to obtain significant information, which can be used to judge the similarity. This can be the position of significant landmarks, or the parameterization of certain shapes within the images, which are obtained by segmentation [10, 11, 12, 13]. Intensity-based measures use the full raw image information. A similarity measure is derived using all intensity values in the two images. However, one may consider introducing the region of interest in order to omit non relevant image parts. Working with this kind of measure is often referred to as voxel property based registration too. The main advantage is that registration can be executed right after image acquisition and definition of an initial pose [10, 12, 13]. Intensity-based measures compare the intensity values of both images pair-wise at the same pixel positions. Subsequently one...
single value is composed out of it with a certain scheme. An advantage of this kind of measure is that it can be used not only with 2D images, but with any kind of data in arbitrary dimensions, as no spatial information is considered. Following measures are included in this group: sum of squared difference (SSD), sum of absolute difference (SAD) and normalized cross correlation (NCC).

$$NCC = \frac{\sum_{i,j \in T} [I_1(i,j)-\bar{I}_1][I_2(i,j)-\bar{I}_2]}{\sqrt{\sum_{i,j \in T} [I_1(i,j)-\bar{I}_1]^2} \cdot \sqrt{\sum_{i,j \in T} [I_2(i,j)-\bar{I}_2]^2}},$$  \hspace{1cm} (10)

where $T$ is the overlap domain of the images, $\bar{I}_1$ and $\bar{I}_2$ denote the mean intensity values in the images $I_1$ and $I_2$ respectively. That means for practical use, that both different contrast and brightness values in the images should not affect the similarity measure, which is very desirable [12, 13, 14]. Disadvantage of the NCC is that the image data has to be read twice: once for calculating the mean intensities and once for completing the summation.

By using horizontal (11) and vertical (12) Sobel templates, four gradient images $dI_1/di$, $dI_2/di$, $dI_1/dj$ and $dI_2/dj$ are created. Then, NCC is calculated of the horizontal and vertical gradient images, which are referred to as $NCC_{difH}$ and $NCC_{difV}$, respectively. The final value $NCC_{dif}$ is the average of the both measures.

$$H = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \hspace{1cm} V = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$ \hspace{1cm} (11 \hspace{0.5cm} 12)

Gradient difference (GD), pattern intensity (PI), gradient correlation (GC) and sum of local normalized correlation (SLNC) are included in group of the measures based on spatial information. A very important advantage of the gradient measure is a fact that they filter out low spatial frequency differences between two images (e.g. caused by soft-tissue structure) and concentrate the contributions to the similarity measure on edge information [13, 14]. GD (5) evaluates two difference images $I_{difH}$ and $I_{difV}$, from gradient images are calculated.

$$GD = \frac{\sum_{i,j \in T} A_H}{A_H + \left[I_{difH}(i,j)\right]^2} + \frac{\sum_{i,j \in T} A_V}{A_V + \left[I_{difV}(i,j)\right]^2},$$ \hspace{1cm} (13)

where $A_H$ and $A_V$ are constants, which denote variance of the respective reference image.

**5 Numerical Results**

In the fig. 3 and 4 are shown entropy ($k=3$) and energy ($k=2$) for T1- and T2- weighted MR.

![Fig. 3 Entropy measure of fuzziness of slices in fig. 2 (a) T1-weighted (b) T2-weighted](image-url)
knee images measure of fuzziness for the original images given in the fig. 2. The fig. 5 illustrates a process of choosing two corresponding slices from the T1- and T2-weighted groups. In these characteristics are depicted correlation between slice number and similarity measures ($G_{D_{norm}}$, $N_{C_{norm}}$, $N_{C_{diff}}$, $N_{C_{normE(A)}}$, $N_{C_{normH(A)}}$). Process of choosing the most similar pair of slices relied on a comparison each image belong to T1-weighted slices group with each image belong to T2-weighted slices group. Application of gradient measures $G_{D}$ and $N_{C_{diff}}$ give a correct registration, but (especially for $G_{D}$) small error caused noisy image can give an incorrect result in registration process. When we calculate in the first step entropy and energy measure of fuzziness for the original images and next use NCC measure, we will obtain a quadruple improvement of the recognition effectiveness. In this case a small noise or distortion do not affect on a correct registration process (in experiment were tested: salt and pepper noise and Gaussian white noise). In the Tab. 1 are shown boundary values of the noise density for salt and pepper noise and border values variance for Gaussian white noise (mean value $M=0$).

### Table 1  Boundary values of the noise

<table>
<thead>
<tr>
<th>Similarity measures</th>
<th>Salt and pepper noise ($D$ is the noise density)</th>
<th>Gaussian white noise (mean $M=0$ and variance $V$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SNR</td>
<td>SNR</td>
</tr>
<tr>
<td>$G_{D}$</td>
<td>0.00235</td>
<td>0.00135</td>
</tr>
<tr>
<td>$N_{C}$</td>
<td>0.00425</td>
<td>0.00475</td>
</tr>
<tr>
<td>$N_{C_{diff}}$</td>
<td>0.00425</td>
<td>0.00875</td>
</tr>
</tbody>
</table>

![Fig. 5 Similarity measures of (a) slice no. 7 and (b) slice no. 15](image-url)
In this experiment parameters and functions were selected as follows: for energy measure of fuzziness $\lambda=0$, $f(x)=x^2$, $F_2(z)=z^2$ and $dv=dx$, for entropy measure of fuzziness $\lambda=0$, $h(x)=x^2$, $F_1(z)=z^2$ and $dv=dx$. The methodology has been tested on 58 clinical T1- and T2-weighted MR knee studies. In all cases registration process has been performed correctly. In my opinion the proposed method of finding corresponding slices seems to be very effective and promising in the registration process.

6 References

Use of data mining methods to improve sleep/wake stages classification

Lukáš Zoubek¹², Sylvie Charbonnier², Suzanne Lesecq²
¹ VSB-Technical University of Ostrava, Department of Measurement and Control, Czech Republic
² Laboratoire d’Automatique de Grenoble, Ecole Nationale Supérieure d’Ingénieurs Electriciens de Grenoble, France
lukas.zoubek@vsb.cz, Sylvie.Charbonnier@inpg.fr

Abstract: This paper describes the process of sleep/wake stage classification and focuses mainly on selection of relevant features extracted from the polysomnographic recordings – EEG, EOG and EMG. Iterative features selection methods were used to improve the classification based on the features characterizing relative power of EEG in five frequency bands. The tests were done on a large database composed of 47 night recordings from different healthy adults.

1 Introduction

Polysomnography consists in the analysis of electroencephalogram (EEG), electromyogram (EMG) and electrooculogic (EOG) signals to recognize the different sleep/wake stages: wake, NREM sleep stages 1 to 4 and Paradoxical Sleep. Recognition of sleep stages from physiological signals recorded during a night sleep enables to build a hypnogram, a temporal succession of sleep/wake stages. Hypnogram is a tool used to diagnose sleep disorders. It is an overall representation of the sleep architecture. Until now, the analysis of a night sleep has been made visually by the physician, who scores every 20s of the recording (named epoch) into one of 6 sleep/wake stages (wake, NREM I, NREM II, NREM III, NREM IV, Paradoxical Sleep). This manual classification is a tedious and time-consuming task. Thanks to the development of computer technology, automated systems to built hypnograms have emerged, either using classical algorithms or artificial intelligence methods, such as neural networks [1], [2]. Features used for classification are extracted from each epoch using signal processing techniques operating in the time domain or in the frequency domain. Studies are still in progress to improve the performance of automatic sleep/wake classifiers. This paper focuses on application of data mining methods used to extract knowledge about sleep/wake classification. The outline of the paper is the following. The whole database and the features extracted are presented in the second section. Features selection methods and classifiers are described in section 3. Results obtained are presented and discussed in section 4.

2 Materials

In this study, a large database of polysomnographic recordings has been used. The full database contains 47 whole night recordings from 41 healthy subjects (19 – 47 years old, 39 males and 2 females). Recordings have been made continuously during the whole night. Four EEG channels (C3-A2, P3-A2, C4-A1, and P4-A1), one diagonal electro-oculogram (EOG) and one chin electromyogram (EMG) have been registered and digitized with the sampling frequency $f_s = 128$ Hz. The EEG signals were measured on the scalp according to the International 10-20 EEG System of Electrodes Placement.
All the 47 polysomnographic recordings were separately and visually classified by two experts. Visual classification was performed on each epoch of the whole night sleep according to the classical sleep stage classification manual [3]. Each epoch was classified into one of 5 different stages (wake, NREM I, NREM II, NREM III&IV, and Paradoxical Sleep) defined for sleep stage classification. Only the epochs classified in the same stage by both experts have been considered in this project. They represent 84% of the whole recordings and form the database. The first line of Table 1 presents the number of epochs classified in each sleep stage for the database. The total number of epochs is 63,254. As could be seen in the Table 1, during a night sleep, the number of epochs classified in a stage is not the same for every stage. Stage NREM II lasts a long time, whereas stage NREM I is rather short. In this study, the database has been reduced to a smaller one where each class is composed of about the same number of epochs, so as to avoid errors in the classification results that could be induced by this difference. The numbers of epochs classified in each sleep stage for the reduced database are presented in the second line of Table 1. The database used in this study consists then of 10,000 randomly selected epochs, each classified into one of the five sleep stages by both experts. This set \( S \) of 10,000 classified epochs is then split in ten subsets \( S = \{S_1, S_2, \ldots, S_{10}\} \), each subset \( S_i \) contains 1,000 epochs, equally distributed in the five classes.

**Table 1 Description of the database used in this study. Number of epochs in the sleep stages.**

<table>
<thead>
<tr>
<th></th>
<th>wake</th>
<th>NREM I</th>
<th>NREM II</th>
<th>NREM III&amp;IV</th>
<th>REM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full database</td>
<td>5,232</td>
<td>1,989</td>
<td>32,966</td>
<td>7,701</td>
<td>15,366</td>
</tr>
<tr>
<td>Test database</td>
<td>1,914</td>
<td>1,879</td>
<td>2,206</td>
<td>1,902</td>
<td>2,099</td>
</tr>
</tbody>
</table>

### 2.1 Features extracted from the physiological signals

Each epoch stored in the database consists of a 20 seconds recording of six signals (four EEG, one EOG and one EMG). During these 20 seconds, the physiological signals are assumed to be stationary. Features are extracted from each epoch using the 6 signals. They are extracted using different signal processing techniques. The PRANA software was used to visualize the polysomnographic recordings and to extract the features. In order to reduce the influence of extreme values that are often observed on physiological variables, each feature of the database was first transformed using nonlinear transformations (transformations towards normal distribution) [4]. They were then normalised using z-score normalisation. Each epoch is represented by a set of 19 features, which are summarized in Table 2.

#### 2.1.1 EEG features

- Five features express the relative power of the EEG signal in given frequency bands. They are calculated using Fourier transformation. The defined frequency bands are: \( \delta_{FT} (0.5 – 4.5 \text{ Hz}), \theta_{FT} (4.5 – 8.5 \text{ Hz}), \alpha_{FT} (8.5 – 11.5), \sigma_{FT} (11.5 – 15.5 \text{ Hz}) \) and \( \beta_{FT} (15.5 – 32.5 \text{Hz}) \). Total spectral power (Ptot) is computed in frequency band (0.5 – 32.5 Hz).

- Five features characterize the wavelet coefficients generated by discrete wavelet transformation. In this study, a 4-level wavelet packet decomposition of the EEG signal has been used to construct a set of wavelet coefficients. The signal decomposition has been performed using the Daubechies3 wavelet. Only the coefficients containing frequency information about the five frequency bands \( \delta_{WT} (0 – 4 \text{ Hz}), \theta_{WT} (4 – 8 \text{ Hz}), \alpha_{WT} (8 – 12 \text{ Hz}), \sigma_{WT} (12 – 16 \text{ Hz}) \) and \( \beta_{WT} (16 – 32 \text{ Hz}) \) have been considered. The information contained in the selected arrays of wavelet coefficients is characterized by the quadratic mean value (root mean square value, RMS) of the coefficients:
\[ \text{RMS}_{FB} = \sqrt{\frac{1}{m-1} \sum_{i=1}^{m} c_{FB}(i)^2} \]  

where \( m \) is the number of wavelet coefficients \( c_{FB}(i) \) in each frequency band \( FB \) and \( FB \in \{ \delta_{WT}, \theta_{WT}, \alpha_{WT}, \beta_{WT} \} \). The features are then expressed as relative values of \( \text{RMS}_{FB} \) computed over these five frequency bands and are labeled as \( \{ \text{RMS}_{rel} \delta, \text{RMS}_{rel} \theta, \text{RMS}_{rel} \alpha, \text{RMS}_{rel} \sigma, \text{RMS}_{rel} \beta \} \).

- Three features describe the signal in the time domain, i.e. the standard deviation, the skewness and kurtosis numbers.

## 2.1.2 EMG and EOG features

EMG and EOG signals have been processed in the time domain. Both signals are characterized by their standard deviation \( \{ \text{std}_{EMG}, \text{std}_{EOG} \} \), their skewness \( \{ \text{skew}_{EMG}, \text{skew}_{EOG} \} \) and their kurtosis \( \{ \text{kurt}_{EMG}, \text{kurt}_{EOG} \} \).

### Table 2 Set of features used in the study to characterize an epoch

<table>
<thead>
<tr>
<th>Feature</th>
<th>EEG signal</th>
<th>EEG signal</th>
<th>EEG signal</th>
<th>EMG signal</th>
<th>EOG signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_{rel} \delta )</td>
<td>( \text{RMS}_{rel} \delta )</td>
<td>( \text{std}_{EEG} )</td>
<td>( \text{skew}_{EEG} )</td>
<td>( \text{kurt}_{EEG} )</td>
<td>( \text{std}_{EOG} )</td>
</tr>
</tbody>
</table>

## 3 Feature selection methods

In this section, the methods used to select the most relevant features are presented. Let \( f_1, f_2, \ldots, f_n \) be a set of \( n \) features to select. Let \( F \) be a subset of these \( n \) features and \( F \) be the subset of features that are not in \( F \): \( F \cup \overline{F} = \{ f_1, f_2, \ldots, f_n \} \) and \( F \cap \overline{F} = \emptyset \). Let \( J \) be a criterion to be maximised and \( J(F) \), the criterion \( J \) that is calculated with the features contained in the subset \( F \). The sequential selection is an iterative technique which selects at each step \( i \) the subset \( F_i \) of features that maximises \( J \) defined by the equation (3).

### 3.1 Sequential Forward Selection (SFS)

The method consists in increasing at each step \( i \) the number of features contained in \( F_{i-1} \) by one. Let \( F_{i-1} \) be the subset of features selected at step \( i-1 \), that maximises \( J \). \( F_{i-1} \) contains \( i-1 \) features, which were previously selected. \( F_{i-1} \) contains the \( n-i+1 \) features still to be selected. At step \( i \), a new feature \( f_i \) is selected out of \( F_{i-1} \) as \( J(F_{i-1} \oplus f_i) = \max(J(F_{i-1} \oplus f_k) \) with \( f_k \in F_{i-1} \). The first subset is initialised to the empty set \( F_0 = \{ \emptyset \} \).

### 3.2 Sequential Backward Selection (SBS)

It consists in decreasing at each step \( i \) the number of features contained in \( F_{i-1} \) by one. Let \( F_{i-1} \) be the subset of features selected at step \( i-1 \), that maximises \( J \). \( F_{i-1} \) contains \( n-i+1 \) features, which were previously selected. \( F_{i-1} \) contains the \( i-1 \) features that have been rejected. At step \( i \), a new feature \( f_i \) is rejected out of \( F_{i-1} \) as \( J(F_{i-1} - f_i) = \max(J(F_{i-1} - f_k) \) with \( f_k \in F_{i-1} \). The first subset is initialised to the subset containing all the features. \( F_0 = \{ f_1, f_2, \ldots, f_n \} \).
3.3 Criterion

In this study, the criterion \( J \) to be maximised is a function of the percentage of epochs correctly classified by a classifier \( C \). As presented in section 2, the data base \( S \) has been split into 10 subsets, \( S = \{ S_1, S_2, ..., S_{10} \} \). Each subset \( S_k \) contains 1,000 epochs, where each of the five classes is equally represented. A classifier \( C \) is trained on one subset \( S_k \) and validated on the 9 other subsets \( S_{\bar{k}} \), \( S_{\bar{k}} = S - S_k \). An accuracy function is calculated on each of the 9 subsets \( S_{\bar{k}} \) as:

\[
Acc(k, \bar{k}) = \frac{\text{card} \left( \{ \text{epoch}(i) \in S_{\bar{k}} / C(\text{epoch}(i)) = E(\text{epoch}(i)) = 0 \} \right)}{\text{card} \left( S_{\bar{k}} \right)}
\]

where \( \text{epoch}(i) \) is an epoch belonging to \( S_{\bar{k}} \), \( C(\text{epoch}(i)) \) is the class assigned to \( \text{epoch}(i) \) by the classifier \( C \), trained on the subset \( k \). \( E(\text{epoch}(i)) \) is the class assigned by the experts to \( \text{epoch}(i) \). A circular permutation is performed on the 10 subsets \( S_k \). The classifier is trained 10 times using the different data sets \( S_k \). Thus, 90 values of \( Acc(k, \bar{k}) \) are obtained. The criterion \( J \) used to select the features is:

\[
J = \frac{1}{10} \sum_{k=1}^{10} \left( \frac{1}{9} \sum_{j \neq k} Acc(k, j) \right)
\]

\( J(F_i) \) is the value of criterion \( J \) defined by (2) and (3) using the features contained in the feature subset \( F_i \).

In equation (3), the term in brackets corresponds to the mean accuracy obtained on the 9 validation sets, when the classifier \( C \) is trained on one training set. \( J \) corresponds to the mean accuracy obtained on the validation sets, when the classifier \( C \) is trained 10 times with 10 different training sets. Calculating \( J \) this way ensures that the accuracy obtained is insensitive to the used training set. The standard deviation of the accuracy \( Acc \) obtained using classifier \( C \) can be computed by equation (4) and can indicate the dispersion of the results.

\[
\text{std}_{Acc} = \left[ \frac{1}{99} \sum_{k=1}^{10} \left( \sum_{j \neq k} (Acc(k, j) - J)^2 \right) \right]^{\frac{1}{2}}
\]

4 Results

The results obtained by the data mining methods are presented in this section. The features selection methods have been processed with an automatic classifier based on the neural networks [2], [5]. In this study was used a multi layer perceptron (MLP) classifier. The architecture of the neural network is the same as in [4]. The results presented below are coming from the C3-A2 EEG channel. Tests have been performed using each of the four EEG channels. The results show that no EEG channel outperforms the others. The mean accuracies obtained using criterion \( J \), defined by (3), were not statically different, whatever the channel used.

4.1 Comparison of Fourier and Wavelet transform

In this first part, the Wavelet decomposition ability to process the EEG signal is analysed. To do so, only the features describing the frequency bands of EEG signals, using the Fourier transform or the Wavelet transform have been used to train the classifiers. The classification
accuracy (3), and its standard deviation (4), obtained using only the features \{P_{rel}^{\delta}, P_{rel}^{\theta}, P_{rel}^{\alpha}, P_{rel}^{\sigma}, P_{rel}^{\beta}\} extracted from the Fourier transform or using only the features obtained thanks to the Wavelet transform \{\text{RMS}_{rel}^{\delta}, \text{RMS}_{rel}^{\theta}, \text{RMS}_{rel}^{\alpha}, \text{RMS}_{rel}^{\sigma}, \text{RMS}_{rel}^{\beta}\} have been computed. Table 3 shows that the accuracy is higher (t-test; p = 0.01) when the relative power in the frequency bands are calculated from the Fourier transform.

Table 3 Classification accuracies for the initial tests – mean values and standard deviations

<table>
<thead>
<tr>
<th>Classification accuracy [%]</th>
<th>Fourier transform</th>
<th>Wavelet transform</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>71.56 ± 1.46</td>
<td>68.85 ± 1.23</td>
</tr>
</tbody>
</table>

4.2 Selection of the most relevant features

The Sequential Forward Selection (SFS) and the Sequential Backward Selection (SBS) methods have been both applied to the set of features presented in Table 2. The features corresponding to EEG signal processed by Wavelet transform (RMS_{rel}^{EEG}) were removed from the set. The subset of features representing the relative power of EEG in the frequency bands obtained thanks to the Fourier transform was considered as a single feature (P_{rel}^{EEG}). The results obtained using the SFS are shown in Fig. 1. The points show the classification accuracy (3), obtained at each step of the process, the bars express the standard deviation (4). The axis of abscissas shows the features selected at each step. The most relevant feature is the set expressing EEG relative power in frequency bands, which is able to correctly classify 71% of the epochs.

The accuracy is significantly increased (t-test; p=0.01) when the kurtosis of EOG, the standard deviation of EEG and the standard deviation of EOG are added (increase from 71.56 ± 1.46% to 75.57 ± 1.80%). Adding the other features, i.e. the features extracted from EMG or the skewness numbers does not improve the classification accuracy. The optimal set of features is then \{(P_{rel}^{\delta}, P_{rel}^{\theta}, P_{rel}^{\alpha}, P_{rel}^{\sigma}, P_{rel}^{\beta}), kurt_{EEG}, std_{EEG}, std_{EOG}\}.

When applying the SBS method, the features from the EMG, the skewness numbers of all the three signals and the kurtosis of the EEG have been detected as non relevant features (there is no decrease in the classification accuracy when these features are removed).

Table 4 Confusion matrixes – classification with EEG relative power in frequency bands (P_{rel}^{EEG}) and with optimal set of features

<table>
<thead>
<tr>
<th>%</th>
<th>wake</th>
<th>NREM I</th>
<th>NREM II</th>
<th>NREM III, IV</th>
<th>PS</th>
</tr>
</thead>
<tbody>
<tr>
<td>wake</td>
<td>81.05</td>
<td>10.16</td>
<td>2.30</td>
<td>1.77</td>
<td>4.72</td>
</tr>
<tr>
<td>NREM I</td>
<td>10.86</td>
<td>45.42</td>
<td>6.19</td>
<td>1.02</td>
<td>36.51</td>
</tr>
<tr>
<td>NREM II</td>
<td>1.66</td>
<td>4.56</td>
<td>79.35</td>
<td>11.14</td>
<td>3.29</td>
</tr>
<tr>
<td>NREM III, IV</td>
<td>0.56</td>
<td>0.42</td>
<td>9.23</td>
<td>88.43</td>
<td>1.36</td>
</tr>
<tr>
<td>PS</td>
<td>5.81</td>
<td>27.65</td>
<td>2.95</td>
<td>0.77</td>
<td>62.82</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>%</th>
<th>wake</th>
<th>NREM I</th>
<th>NREM II</th>
<th>NREM III, IV</th>
<th>PS</th>
</tr>
</thead>
<tbody>
<tr>
<td>wake</td>
<td>84.41</td>
<td>7.90</td>
<td>2.16</td>
<td>1.28</td>
<td>4.25</td>
</tr>
<tr>
<td>NREM I</td>
<td>9.58</td>
<td>55.22</td>
<td>7.20</td>
<td>1.65</td>
<td>26.35</td>
</tr>
<tr>
<td>NREM II</td>
<td>1.63</td>
<td>4.21</td>
<td>82.77</td>
<td>8.96</td>
<td>2.43</td>
</tr>
<tr>
<td>NREM III, IV</td>
<td>1.18</td>
<td>0.27</td>
<td>8.69</td>
<td>89.74</td>
<td>0.12</td>
</tr>
<tr>
<td>PS</td>
<td>6.64</td>
<td>24.09</td>
<td>3.38</td>
<td>0.71</td>
<td>65.18</td>
</tr>
</tbody>
</table>
Table 4 compares the confusion matrixes obtained when only $P_{rel}$ EEG features are used and then when the optimal set of features is used. The columns represent the stages predicted by the classifier and the rows represent the stages determined by the physician. Wake, NREM II and NREM III&IV are correctly classified using EEG relative power spectra, whereas NREM I and PS stages are more difficult to distinguish. Adding the $kurt_{EOG}$, the $std_{EEG}$ and the $std_{EOG}$ improves significantly NREM I accuracy and slightly wake, NREM II and PS accuracies. The analysis of the confusion matrix at each step of the selection, using SFS, shows that the stage NREM I accuracy is especially improved by $kurt_{EOG}$. This feature helps to distinguish stages NREM I and Paradoxical Sleep, which are hardly distinguishable by the EEG signal analyzed in the frequency domain. The parameter $std_{EEG}$ improves the wake and PS stage accuracies.

5 Conclusion

In this paper, data mining methods have been applied on a large database in order to select the most relevant features for sleep/wake stage classification. The results show that appropriate selection of the features improves the classification of sleep stages. The relative power spectrum of EEG, computed over different frequency bands, enables to correctly classify about 71% of the sleep stages. Adding the standard deviation of EOG, the standard deviation of EEG and the kurtosis of EOG improves the classification accuracy by about 4%. Information extracted from the EMG signal appears to be irrelevant for sleep stage classification, but this may due to the artefacts affecting this signal. Future work could be oriented on the extraction of new features from the polysomnographical signals to improve discrimination between stages NREM I and Paradoxical Sleep. Indeed, the classification accuracy of these two stages is lower compared to the other sleep stages. EMG could also be introduced in the decision system, separately from the other signals.

6 Acknowledgement

Special thanks are expressed to Alain Buguet and Florian Chapotot for preparation of the sleep recording database and to the PhiTools (Strasbourg, France) for lending the PRANA software and the sleep recording database. This work was partially supported by the faculty internal project, "Biomedical engineering systems II", IGA BI4556011.

7 References

Section 4

Telecommunication
Stochastic models for multiplexed VoIP traffic

Arkadiusz Biernacki
Silesian University of Technology, Institute of Computer Science
Arkadiusz.Biernacki@polsl.pl

Abstract: The concept of multiplexing computer network traffic on a common channel for efficient utilization of the transmission link capacity is a great concern to network engineers. A statistical multiplexer allocates a channel capacity that lies between the average and peak rates and buffers the traffic during periods when demand exceeds channel capacity. In this work we review existing models for multiplexing of independent packet voice sources, which alternate between bursts of transmission activity and periods of no activity. We study the models taking into account their accuracy, traffic behaviour level (packet, burst, connection). We also propose our own solution for Voice over IP protocol (VoIP) multiplexing, which is based on Markov Process and approximates traffic behaviour in burst and connection scales.

1 Introduction

The growth of communication based on Voice over IP protocol (VoIP) has been exceptional during recent years and is expected to continue in the future. Consequently, voice packets produced during telephone conversations are to have considerable share in all voice packets sent through networks. When certain amount of voice calls is performed simultaneously on a single link, the link needs to be shared between them, and a statistical multiplexing of voice packets is necessary. The device, which handles the multiplexing, is called a multiplexer. Two types of multiplexing are possible: a frequency division multiplexing (FDM), where each user get an exclusive share of the links capacity and a time division multiplexing (TDM), where periodically one user at a time gain control of the full capacity of the link for a short instance of time. In this work we are taking into account the TDM multiplexing.

Statistical multiplexers usually are modelled as queuing systems with buffer space, served by one or more transmission links of fixed or varying capacity. If the sum of the peak rates $P_i$, is not allowed to exceed the output link rate $C$ of the multiplexer, then we say that the multiplexer is working under peak rate allocation. The advantages of peak rate allocation multiplexing are no packet loss due to buffer overflow at burst level as well as minimal packet delay. The disadvantage is that bandwidth is wasted when the input links are sending at a lower rate than their peak rate $P_i$, which happens when multiplexed are variable bit rate (VBR) sources. This motivates the argument for statistical multiplexing where the sum of the connection peak rates is allowed to exceed the link capacity.

The ratio of the number of VBR sources that can be multiplexed on a fixed capacity link under a specified delay or loss constraint and the number of sources that can be supported on the basis of peak rate allocation is called statistical multiplexing gain (SMG). To determine and maximize the SMG, admission control rules are formulated that relate to traffic characteristics, which flow into the buffer of multiplexer, to performance constraints and system parameters. Aforementioned characteristic of the traffic is of important interest amongst network engineers and rest of the paper will concentrate on this subject.

The rest of this paper is organized as follows: in section 2 we gave basics concepts of VoIP traffic modelling. In section 3 we described the previous researches in the area of packetized
voice multiplexing. In section 4 we presented our model of multiplexed VoIP traffic and validated it in section 5. Section 6 is a conclusion of our work.

2 Key concepts

2.1 Markov process

Markov processes provide very flexible, powerful, and efficient means for the description and analysis of dynamic (computer) system properties and can be used to describe a system moving over time between different states.

Definition 1 A stochastic process is defined as a family of random variables \( \{ X(t) : t \in T \} \) where each random variable \( X(t) \) has a \( t \) parameter, which can be called the time parameter.

Definition 2 A stochastic process \( \{ X(t) : t \in T \} \) constitutes a Markov Process (MP) if for all \( 0 = t_0 < t_1 < \ldots < t_n < t_{n+1} \) the conditional cumulative distribution function of \( X(t) \) depends only on the last previous value \( X(t_n) \) and not on the earlier values \( X(t_{n-1}), \ldots, X(t_0) \):

\[
P[X(t) \leq x | X(t_n) = x_n, X(t_{n-1}) = x_{n-1}, \ldots, X(t_0) = x_0] = P[X(t) \leq x | X(t_n) = x_n]
\]

This most general definition of a Markov process can be adopted to special cases. Our model, presented in section 4, is based on discrete state spaces, \( \forall x_i \in \mathbb{N} \), and continuous time parameter \( t \). Such process is called continuous-time Markov chains (CTMC). The state sojourn times of a CTMC are exponentially distributed.

Definition 3 CTMC is time-homogeneous when \( P[X(t) \leq x | X(t_n) = x_n] \) depends only on \( (t - t_n) \) and is not function of \( t \) and \( t_n \).

For time-homogeneous CTMC, for very small interval \( \tau = \Delta t \), there is linear dependency

\[
p_{ij} = q_{ij} \Delta t , \tag{1}
\]

where \( p_{ij} \) is probability of transition between state \( i \) and \( j \), and \( q_{ij} \) is transition rate (intensity) coefficient [1].

2.2 VoIP traffic characteristic

A VoIP conversation, like traditional telephony conversation, can be considered as an alternate process during which, one of interlocutors is speaking and the other one is listening. Moreover, the speaking interlocutor often makes small gaps between spoken phrases, words and syllables. Aforementioned characteristics of a conversation are used by a voice coder with voice activity detector (VAD), which detects frames containing silence and suspends them from further emission through packet network, fig. 1 left. Besides, voice in a coder is also compressed. The above mentioned operations allow for bandwidth consumption reduction in packet network transporting VoIP packets. The VoIP source can be considered as two state Markov chain, fig 1 right, which is in ON state when interlocutor speaks and is in OFF state when interlocutor remains silent. Statistical distribution of time spend by the source in these two states plays significant role in multiplexed VoIP model creation and was obtained from [2].
3 Review of previous works

The performance analysis of packetized voice traffic usually includes the analysis of the appropriate queuing model. The work related to the analysis can be divided on two groups.

In the first group, authors concentrated on a microscopic view of network traffic and try to model dependency between subsequent packets (micro scale). The model of the superposition range from an exact Markov Renewal Process to an approximate renewal process mode. Eckberg treat multiplexed voice as $\sum D_i / D/1$ and has derived the exact delay distribution for a it [3]. In [4] and [5] it was stated that the multiplexed voice streams may be approximated with quite good results by a Poisson process. In [6] renewal processes were used and voice multiplexer was modelled as $G_i / D/1$ queuing system. Albin [7] based his analysis on $\sum G/1/1$ queuing systems and renewal processes.

The second group consist of works in which authors try to match the behaviour of the point process over a relatively long time interval (macro scale) neglecting dependency between subsequent packets. Usually, the statistical properties of voice source are taking into account. Stern [5] has presented a queuing model based on the exponential ON/OFF model and an imbedded continuous time Markov chain whose states represent the number of currently active speakers. Daigle et. all in [9] investigated three different approximations for aggregate arrival process based on exponential ON/OFF sources has been compared with simulations and the delay performance in a statistical multiplexer has been examined. In [10] a multiplexer with infinite buffer was studied with a Stochastic Fluid Flow model but it is shown in [11] that this model only work for a multiplexer under heavy load. A multiplexer with finite buffer is studied in [12] using the fluid flow model but it does not work well for small buffers.

Some of the authors proposed approximate methods. A two-state Markov Modulated Poisson process (MMPP) is used quite successfully in [13] to estimate the delay in a multiplexer with infinite buffer an it is suggested that the same approach for calculating the parameters of the MMPP can be used for a multiplexer with finite buffer, but in [14] shown that this does not work in the finite buffer case. Instead is a different method for finding the parameters of the MMPP developed. Besides authors proposed two other concept based on renewal processes and fluid models to estimate multiplexer efficiency. In [15] is the arrival process approximated with a two-state MMPP and a method called asymptotic matching is suggested for the calculation of the parameters of the MMPP. The common conclusion of macro scale models is that fluid models lack of stochastic properties of process but is better for correlation modelling in comparison to MMPP.

Our model, which neglects inter-packets dependencies, can be classified to the second group. The first factor that distinguishes our model form from the competition is it is taking into
account a connection scale. The second factor is updated set of parameters for ON/OFF periods, which reflects technical advance in VAD by voice coders.

4 Proposed model

We propose a model based on two-dimensional Markov Process, which acts at two different time scales – connection and burst, which have physical interpretation, fig. 2. Connection scale describes conversations durations and their interarrival times. Connections are generated according to a Poisson process with rate $\lambda_c$. Their durations are exponentially distributed with mean $\mu_c^{-1}$ [16].

During each connection packet flows are generated. Assuming the VoIP source to be two-state ON/OFF model, for each connection, number of flows forms a bird-death process and the maximum flows number depends on current number of connection. The transition rates between flows (eq. 1) are based on $\mu_b$ and $\lambda_b$ transition rates and additionally depends also on the number of connections. By $\alpha$ we denoted probability, that with new connection beginning, new flow is also generated and it equals to probability that the VoIP source is in

Fig. 2. Multiplexed VoIP traffic model based on CTMC
ON state. The number of states in the model is equal to \( \sum_{i=1}^{n} i \) where by \( n \) we denoted number of maximum allowed connections.

5 Model validation

We compared mean, variation and the index of dispersion for counts (IDC) of real multiplexed traffic with the same statistics of our model, which were computed analytically, directly from the CTMC – table 1. The IDC is an indicator of a traffic stream variability and was computed as a quotient of a variance to mean, \( IDC = \frac{Var[X(t)]}{E[X(t)]} \) [17].

Then, we constructed traffic generator based on our model transformed to continuous-time Markov reward chain, where a reward rate is associated to each state in the model. The transmission rate when the model is at state \( j \) is \( r_j \) packets per second. The traffic generator tool obtains the state transition probability between any two states as follows. Let \( \Delta_j \) be the sum of the output rates out of state \( j \), that is, \( \Delta_j = \sum_k \lambda_{j,k} \), where \( \lambda_{j,k} \) is the transition rate from state \( j \) to \( k \). Then the transition probability from state \( j \) to \( k \) is \( P_{jk} = \frac{\lambda_{j,k}}{\Delta_j} \). The number of packet sent in state \( j \) equals \( \Delta_j r_j \).

Taking into account the above formulas, we generated some synthetic traces and compared their autocorrelation with autocorrelation of real traffic – fig. 3.

| Table 1 Comparison between real traffic and the traffic model for max. 60 active lines |
|---------------------------------|----------------|----------------|
|                                | Real traffic  | Traffic model |
| Mean                            | 495           | 450           |
| Variation                       | 18620         | 15500         |
| IDC                             | 37            | 34            |

Fig. 3. Comparison of autocorrelation functions for real and synthetic VoIP traces: left figure – 30 lines, right figure – 60 lines multiplexed

6 Conclusions

In this paper we studied multiplexed VoIP traffic models taking into account their accuracy and traffic behaviour level (packet, burst, connection). We also proposed our own solution for VoIP packets multiplexing, which was based on Markov Process and approximates traffic behaviour in connection and burst scales. The proposed model approximated multiplexed real traffic with fair level of accuracy – table 1, although underestimates its autocorrelation which
is visible on fig. 3. Further research will aim at computing efficiency of multiplexing devices during VoIP packets transmissions.

7 References

Admission control for EDCA in 802.11 wireless networks

Agnieszka Brachman
Silesian University of Technology
agnieszka.brachman@polsl.pl

Abstract: The article briefly presents MAC-layer enhancements proposed in the new-released IEEE 802.11e standard in order to provide QoS at MAC level and problems related with adjusting EDCA medium access function parameters to act well under various network conditions. The role of admission control and its tasks are presented as well as AC methods for EDCA mechanism that can be investigated and furthermore applied in wireless networks to achieve required QoS level.

1 Introduction

IEEE 802.11 [1] wireless networks have become incredibly popular. At many public areas, buildings, airports, users access the Internet using deployed hotspots. In addition users expect the possibility of running applications that require transmission of real-time traffic which leads us to the necessity of providing Quality of Service (QoS) in WLANs. QoS mechanisms are implemented to guarantee timely delivery of data to a particular destination. Each application transmitting data has different needs in terms of latency, guaranteed throughput, loss rate, delay and delay jitter. The QoS support is particularly essential in networks where video, voice or audio is delivered, in order to provide satisfactory level of transmission to every customer. Even in high-speed networks QoS is necessary to protect real-time traffic from ordinary data. It is important to provide capability of the network to support different levels of services not only the best effort services as in current solutions.

For a few years providing QoS in wireless networks remains an open issue. Standard IEEE 802.11e [2] introduces enhancements required in QoS provisioning at MAC level. The proposed improvements are widely discussed in [4, 5, 6, 8, 9, 21]. One of the most important issues of global QoS provisioning is admission control (AC). Admission control [4, 7, 8, 9] determines how bandwidth is allocated to traffic streams with various requirements. Its role is to admit or reject the upcoming connection requests to protect existing streams.

This paper presents current state of research on admission control mechanisms for EDCA medium access function in wireless, homogeneous networks. The paper is organized as follows. Section 2 outlines the main features of medium access functions introduced by IEEE 802.11e standard. In section 3 the main issues and problems of admission control in wireless networks are described. The next section depicts recent research works on admission control in wireless networks for EDCA medium access function introduced by 802.11e standard. The main problems and performance evaluation of described methods is presented. Finally the conclusions, future work and references are presented in section 5 and 6.

2 IEEE 802.11e enhancements

Stations in WLANs can be configured into two modes: “ad hoc” and infrastructure. In “ad hoc” mode, all stations remain within the communication range and are able to directly
transmit data with each other, whereas in infrastructure mode an access point (AP) is required. Stations communicate with each other through the AP.

The original 802.11 MAC protocol specified in [1] introduces two methods for accessing the medium: contention-based Distributed Coordination Function (DCF) and pooling-based Point Coordination Function (PCF) [1, 4, 5, 6, 8, 9, 21]. DCF is the basic medium access function that operates in both “ad hoc” and infrastructure-based network configuration. DCF is based on Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) mechanism for sharing the medium. The station that attempts to send a packet of data, obeying to the rules of CSMA/CA, checks if the medium is idle, before transmitting. Once it senses a busy signal, it must determine individually when to access the medium. In order to reduce the probability that two or more stations begin transmission at the same time, each station maintains backoff counter, which is decreased after the medium has been idle for a minimum interval equal to DIFS value. The station may begin transmitting only if the counter reaches 0. The backoff counter is calculated as a random, multiple number of slots, between 0 and Contention Window (CW). Each node holds CW variable and initially sets its value to CWmin, where CWmin is equal for all the stations. If one station begins its transmission, the others freeze their counters and begin to decrease them as soon as the new DIFS elapses, after detecting the medium being idle. If the transmission fails, the transmitter doubles the value of CW and begins a new backoff procedure. When CW reaches CWmax value, the packet is dropped. After a successful transmission, CW is set to CWmin and a new value of backoff counter is randomized. PCF is an optionally implemented method in wireless devices. It requires the presence of a central coordinator called Point Coordinator (PC); typically the function is performed by access point (AP). PC polls awaiting stations according to Round Robin algorithm until contention free period expires or there is nothing to be sent. Details for cooperation of the two methods are described in [1, 4, 9, 21]. DCF offers “best-effort” services. There is no mechanism to provide short- or long-term fairness in accessing the medium or control delays between frames sent by a station. Also, although the PCF was supposed to provide support for real-time traffic, it doesn’t guarantee any time-constraint delivery. The limitations of the DCF and PCF are discussed in [4, 5, 6, 9, 21].

In the upcoming standard IEEE 802.11e [2], DCF is replaced by Enhanced DCF (EDCF). With EDCF, each station can have up to 4 queues, mapped to different traffic classes. Service differentiation can be achieved by assigning different values of parameters for every queue contending for access. The parameters are: CWmin, CWmax, AIFS which is inter frame space earlier described as DIFS, and TxOpLimit which is maximum time of transmission. PCF is replaced with Hybrid Coordination Function. The description of the method is not presented as it is not further discussed in the paper; it can be found in [2, 5, 9, 10, 21].

The EDCA provides a prioritized QoS service using an independent transmit queue for each traffic class. Traffic belonging to a higher priority class has a higher probability of accessing the medium, thus achieving a higher throughput when competing with lower priority traffic. However, service differentiation can provide service prioritization but not QoS guarantee. Many applications would require absolute bandwidth provision from the network and the resource reservation remains an open and significant issue. The problem is especially apparent when the wireless channel is overloaded. Nevertheless EDCA may work well under various conditions when admission control mechanism is applied and used to adapt EDCA parameters for varying network conditions. Appropriate distribution of parameter’s values to achieve target QoS is a challenging issue.
3 Admission control in WLANs

As mentioned above, QoS can be interpreted as the ability of the network to provide consistent services for multimedia delivery and it is necessary to ensure each application transmitting data, allocated bandwidth in particular interspaces. In order to provide QoS-enabled transmission between two end users, QoS must be supported at every stage of the route.

Several models were defined to achieve required QoS for Layer 3, namely: Integrated Services (IntServ) [7] and Differentiated Services (DiffServ) [3, 7]. The Integrated Services model is based on reservation protocol; it is not scaleable and has not been implemented widely. The Differentiated Services model was developed to differentiate IP traffic. Traffic was classified based on priority. This allows transmitting traffic with similar constraints with similar traffic guarantees across multiple networks.

The main approach for QoS support in WLAN is based on the Differentiated Services. It provides service differentiation through simple mechanisms. Examples are described in [7, 8]. These mechanisms, however, mostly rely on centralized control. Standard 802.11e doesn’t define how to apply proposed enhancements in MAC layer in various scenarios of providing QoS. It doesn’t specify any admission control mechanisms, which should concern customization of MAC parameters, according to which algorithm acts.

The main role of admission control is to limit traffic in particular service class (category access), so that the existing flows will not be harmed, and to provide the maximum possible channel utilization. An application that wishes to transmit data with QoS requirements must first request a connection providing information about characteristics of the traffic and the QoS parameters required by the application. The network must decide whether enough resources are available to accept the connection considering previous contracts and then either accept or reject the connection request. This mechanism is known as Admission Control.

Note that to provide QoS, devices must be able to support both: layer 2 prioritization or parameterization (HCF) and layer 3 prioritization (e.g. DiffServ). Providing differentiated services in wireless environment requires that MAC layer will support effective differentiation between different types of services. DiffServ is not a complete answer to all QoS issues nevertheless it is a critical step in achieving this long-term objective.

4 Admission control mechanisms for EDCA in WLANs

Several methods how to apply admission control using 802.11e MAC enhancements were introduced. They are divided into two categories: measured and model based.

4.1 Measured-based admission control methods

4.1.1 Distributed AC

Distributed admission control (DAC) isn’t any method of customizing MAC parameters. This mechanism only protect existing flows [11, 12]. It requires the existence of Point Coordinator (PC). The PC defines transmission limit for every access category (AC). It measures the amount of time occupied by the transmission of traffic in each AC. Basing on these parameters it calculates the transmission budget for each AC and announces it during beacons intervals. Each station receives the information and determines if new transmission can start.
The DAC method is a very basic technique of protecting and guaranteeing existing flows when the traffic load is not very heavy. It doesn’t provide any QoS requirements for an application. It also requires the presence of PC and causes vibration in the network because every station must adjust its parameters at every beacon interval. The overall throughput also decreases under heavy traffic conditions. Similar solution was proposed in [12]. It proposes a scheme in which real-time traffic and data transmission are handled differently. For real-time traffic, DAC algorithm is applied. For data traffic, the PC controls parameters for stations globally based on traffic conditions. The method improves protection offered for video and voice streams via lower throughput for data traffic. It allows better differentiation between different AC and provides fairness for traffic with similar requirements.

4.1.2 Two-level protection and Guarantee Mechanism
This method is based on DAC scheme. It is meant to better protect existing flows in heavier loaded medium [13]. Basing on information provided by PC concerning transmission budget, the new flow is accepted only tentatively and the station measures throughput and delays performance. If any of these parameters exceeds defined threshold the newly accepted stream will kill or reject itself. The new flows will be rejected at the beginning if the transmission budget exceeds established value. This is called the first-level protection and it acts well for protecting each existing QoS flows from new and other existing flows. The second-level protection decreases the amount of best-effort traffic by increasing the initial contention window size and inter-frame space when the number of active QoS flows is large. It provides lower number of collisions. The problem of this solution is that it also doesn’t provide any QoS requirements for an application and it requires optimized values of thresholds.

4.1.3 Virtual MAC and Virtual Source Algorithms
The idea of the solution is to virtually run the application and perform transmission to calculate probable delays and collision rate [14, 15]. The Virtual MAC method consists of handling virtual packets at MAC level i.e. estimating time of sending packet and collision detection afterwards as well as a backoff procedure. The Virtual Source method is performed to count delay for packets generated by virtual application at certain rate to the moment when they are served by VMAC. The main advantage of the solution is that it is performed virtually, so it doesn’t put any additional load to the medium. However it requires additional processing at every station. The decision to the admission of the new flow is taken based on the calculation of delay and collision rate which also may be very fuzzy. The algorithms were proposed before standard IEEE 802.11e was introduced as an essential component of QoS-aware MAC layer. The global solution requires differentiating between real-time traffic and data transmission. The virtual algorithms were performed only for more sensitive traffic streams, and were designed to continuously monitor the channel’s state. As a result this method results in degradation of low priority traffic. It can be used in both “ad hoc” and infrastructure mode. However, since it is virtual, it only considers the effect of existing flows on the incoming flow, not the effects of the incoming flow on existing flows, which may introduce inaccuracy in making admission decisions.

4.1.4 Threshold-based AC
This solution involves that every station measures the traffic condition on the medium [16]. Depending on preferred solution the stations measure relative time when the wireless medium is busy – relative occupied bandwidth \( B_{\text{occu}} \) (Eq. 1) or average collision ratio \( R_c \) (Eq. 2).
\[ B_{\text{occu}} = \frac{T_{\text{Busy}}}{T} \times 100\% \] (1)

where, \( T_{\text{Busy}} \) - amount of time when the medium is busy, \( T \) - sampling period.

\[ R_c = \frac{N_c}{N_t} \] (2)

where, \( N_c \) - number of collision, \( N_t \) - total number of transmissions.

If the measured value is below the lower threshold, the new flow with the highest AC priority is admitted. If it is within lower and upper thresholds, no action is taken. If it exceeds the upper threshold, the transmission of the lowest active AC during the next period should be stopped. A station can easily compute amount of time when the medium is busy using NAV vector, also in 802.11e the station maintains a number of retransmission, which is a good indicator of collision rate. This method is very easy to implement but requires the threshold values, which are difficult to fix. Moreover, they should be changed in order to provide the best performance during changing conditions on the medium.

4.1.5 HARMONICA

The solution was proposed and described in [17]. It requires the presence of access point (AP), which performs the following actions:

- LQI Monitor – it samples link layer quality indicator (LQI), which parameters include drop rate, link layer end-to-end delay, throughput for each traffic class;
- Adaptation Control Unit – it executes two adaptation algorithms based on LQI, to obtain best parameters for each AC;
- Admission Controller – accepts or rejects upcoming flows.

Direct mapping between LQI and QoS requirements is easy to perform at the application layer. The mentioned adaptation algorithms are relative and base adaptation. The first one selects the best differences between parameters for each AC, so that resources can be optimally allocated. The second algorithm increases or decreases category access parameters to achieve high channel utilization. To avoid congestions admission control is performed for each real-time traffic stream. If expected limits on drop rate or delay bound cannot be guaranteed, and there is no possibility of decreasing bandwidth admitted for best-effort traffic class in order to use it for candidate flow, then the flow should be rejected. Best-effort traffic can always be admitted. The solution provides possibility to guarantee a minimal bandwidth level for data categorized to best-effort traffic class. Through dynamically adjusting parameters for each AC it is possible to match QoS requirements and efficiently use network resources. The algorithm doesn’t perform well under light load because the base adaptation algorithm is misled. However under such conditions optimization is less or unimportant. Broadcasting new values of parameters for ACs doesn’t require additional communication while beacon frame introduced in IEEE 802.11e contains defined QoS parameter set.

4.1.6 Assured Rate MAC Extension

The solution is a very simple example of admission control. It is an extension of assured rate service in Differentiated Services for 802.11 WLANs [17]. It allows achieving requested throughput by adjusting contention window size based on observation of the estimated sending and desired rate. It was proposed for legacy IEEE 802.11 standard and it doesn’t take into account the EDCA function. It allows keeping MAC protocol fully distributed. For stations that support ARME method it reduces the size of contention window to achieve higher transmission rate than best effort services. It doesn’t lead to starvation of best effort terminals because it trades off the bandwidth assurance of Assured Rate in case of overload.
4.2 Model-based admission control methods

4.2.1 Markov Chain
This method [18] allows estimating throughput that the entire traffic will achieve after admitting a new flow. The estimations are calculated using collision statistics for each AC. The decision on rejecting or accepting the stream is made based on the calculations. The analytical model is described in [18, 19]. The algorithm considers dynamic network conditions, namely the number of active flows as well as the parameters set for each AC. It estimates the best values for AC parameters under changing network conditions. By avoiding network congestion the existing QoS-enabled flows are protected from incoming transmissions’ requests. The algorithm provides efficient network utilization as well. It works in infrastructure mode. The additional tasks performed by AP are: collision monitoring, throughput estimation, taking admission decisions. The main disadvantage of this algorithm is that it assumes that every station has always data to transmit – the model is derived under saturation conditions. It doesn’t consider the existence of more than one flow at each station. It adjusts only the size of contention window what may lead to longer delays.

4.2.2 Throughput Guarantee Service
The main idea of this solution is to adjust contention window size to provide desired throughput guarantees for accepted flows [20]. When a new flow is attempting to access the network, AP calculates the new values of contention windows size and afterwards probable throughput for each flow. If it meets the requirements, the new flow is admitted and the changes are propagated. As the previous method, this solution has similar problems. It assumes that each station has data to send and that the number of the accepted stations almost saturate the channel. It only adjusts CW parameter which doesn’t fully utilize the improvements introduced by IEEE 802.11e, the value of rest parameters remains fixed as well as AIFS which is equal to DIFS. This doesn’t support best-effort traffic and coexistence of the two types of traffic: best effort and throughput guarantee is very poor. Best effort traffic consumes the leftover bandwidth and has insignificant impact on the guaranteed flows. The algorithm requires presence of AP that acts as a Wireless LAN Bandwidth Broker.

5 Conclusions and future work
EDCA provides prioritized QoS support based on Access Categories (AC) by differentiating inter-frame space, minimum and maximum contention window size and transmission duration time. The method is easier to implement than HCCA and works also in “ad hoc” mode. Simulation results show that increased efficiency using EDCA is achieved, however it is provided at the cost of worst performance for lower priority traffic; also, effective EDCA parameters adjustment algorithm is required. Survey of admission control methods for EDCA is presented. Measurement-based methods are easier to implement, however they require theoretical foundation and sometimes it is very difficult to achieve optimization. Model-based schemes require analytical models derived sometimes on unrealistic assumptions thus the solution may not be suitable for practical systems.

Future work concerns building a model of proposed enhancements and evaluating performance for different scenario to answer the problem how to adapt the EDCA parameters to the varying traffic load and channel condition, to assure better channel utilization and efficiency. The next task is to compare different scheduling schemes for HCCA mechanism, using different scenarios and traffic flows. Afterwards complete and comprehensive solution
for both EDCA and HCCA should be proposed. A richer set of test scenarios should represent a wide variety of situations to make simulation results more realistic.

6 References

ZigBee in Biotelemetry

Floder Jan, Černý Martin
Department of Measurement and Control, VŠB Technical University of Ostrava, Faculty of Electrical Engineering and Informatics
jan.floder@vsb.cz, martin.cerny@vsb.cz

Abstract: In frame of TeleCare and HomeCare was designed and created measuring chain for biosignal measuring and wireless transmission. ZigBee was used for wireless data transmission. Data was than visualized and recorded for further analysis.

In this paper is shown how whole chain and visualization were created. What type of networking was used and is suitable for similar cases, what method of data transmission protocol use, what protocol use or create/modify and how to joint transmission speed request, minimal data loss and minimal energy consumption.

1 Project

In frame of TeleCare project was created low distance wireless network of biomedical measuring devices. The main goal of project is measuring and transfer values about vehicle drivers condition and some additional values about ride. These drivers’ signals are collected: pulse, ECG, SpO2 and body temperature. Besides these signals there are measured values of acceleration of vehicle in X and Y axis (front-back, left-right) and air temperature.

Measuring chain (see Fig. 1) consists of two biomedical measuring devices (ECG, ChipOx.), acceleration data sensing device and PC. Measured biomedical values are transmitted by ZigBee technology to retransmitting device. This retransmitting device collect received values and, in addition, measures values of vehicle acceleration. All data collected by this device are sent, by using radio-modules, to PC. There are all values displayed and saved for further analysis.

![Fig. 1 Designed measuring chain](image-url)
2 ZigBee

ZigBee is a new standard for wireless data transmission. It is an 802.15.4 WPAN standard which was specified in 2003. The first generation of ZigBee standard-compatible products came on the market in 2005. ZigBee is designed for low distance data transmission at lower speeds. It is good for data transmission from sensors and to some control peripherals. Next very big advantage of ZigBee is low energy consumption. It is the reason why most of applications with ZigBee are suitable for battery power.

PAN4551 (from Panasonic) ZigBee modules were used in this project. PAN4551 consists of a transceiver MC13193 and MCU HC9S08GT60 both from Freescale. Transceiver operates in the 2.4GHz band, which is divided into 16 channels. Communication is done in packet or stream mode (both with CRC checking in HW). Maximal transmission speed (between two nodes) is 250 kbps (kilobits). Speed degrades when obstacles are between sender and receiver. Maximal distance between two nodes is tenths meters (when using suitable mini-antennas). Power consumption of chosen PAN4551 is at about 30 ÷ 35 mA during transmission (receive or send). Power consumption can fall down when transceiver is turned to idle state on time, when it isn’t necessary to transmit or receive any data.

3 ZigBee in Project

Designed ZigBee network consists of two slave units (Slave 1 and Slave 2) and one master unit (Master).

![Diagram of designed ZigBee network]

The master unit polls each slave for data from peripheries. In addition the master unit measures values of the accelerometer. It is possible because each PAN4551 module includes a 12-bit, 8-channel A/D converter. After that all the data are collected by the master and sent to the PC using radiomodems.

This wireless network (star topology and master-multislave medium access) is most suitable for our and similar tasks, when all sensors or other peripherals are in range of central node, fast data transmission is needed and slaves do not need to transfer too much data between each other. Everything is controlled by the master, “way” of signal is known and never change (master-multislave medium access). It is easy to reprogram it for different data loads. In case of some nodes aren’t visible, it is possible to use tree net topolgy or in more simplistic cases star, where main node (which provides connection to some data acquisition or control system)

---

1  Maximal measured achievable speed in packet mode (speed depends on packet size) and master–slave configuration (acknowledging) with packet manipulation was at about 10 kbps (kiloBytes).
is not the master unit. Second solution has advantage of one “master” which controls all communication.

![Fig. 3 Changes in topology for more complicated systems](image)

Maximum speed for point-to-point (and master-slave configuration with packet manipulation) communication in packet mode is at about 10 kBps (for one channel). It is good to set up for each branch of tree other transmission channel to minimize possible collisions in networks with more nodes.

### 3.1 Protocol and communication

#### 3.1.1 Protocol

For MC13193 there is no multi-point protocol defined in hardware so user has to create own software protocol or use existing commercial one. Except paid protocols (Z-Stack, etc.) also exist free ones like SimpleMAC (software library) or 802.15.4 standard (bigger and more complicated than SimpleMAC). The SimpleMAC is the simplest one, because compiled it occupies at about 2÷4 kB of flash memory on HCS08. It provides all necessary functions to communicate with transceiver MC13193 and all functions for point-to-point communication between modules in packet mode too. For most complicated purposes, when mesh nets are necessary, it is better to use pay commercial solution (Z-Stack) then create own protocol. SimpleMAC was chosen for project. SimpleMAC library contains only functions for communication with MC13193, sends and receives data in packet mode. It is really small and easy to work with it. It gives us a full control over communication with other nodes and over speed of communication. For non point-to-point communication was created simple protocol above SimpleMAC to simplify communication. This can be useful for nets in star topology. This protocol enables us to address up to 255 nodes, use subnets, commands for sending receiving data, replies on requests, auto/manual data/command acknowledge mode, waiting for specific data. It is highly configurable and can be easily extended. Final solution is very small, it occupies only 1.5 kB of space.

#### 3.1.2 Communication

Communication is done as master-multislave. The master unit polls slave units and requires data packets. Slave units send data or reply that no new data is available or send error code. Except data requests, master sends also commands to initialize measuring devices and for start and stop measurement.

All measured values are buffered for 0.4 second to avoid possible data loss.
The ChipOx sends at about 200 Bps, the ECG 100 sends at about 100 Bps. Sent data size is about 300 Bps, what is far below border of 10 kBps and it is because why is possible to allow delays in communication. In this time is transmitter turned to idle mode to save power. Master collects data from these two slaves and adds data from accelerometer from X and Y axis with sampling rate 100Hz or 50Hz. Amount of data send from master to PC via radio-modem is maximally at about 500Bps.

### 3.2 Programming

PAN4551 is equipped with HC9S08GT60 with 60kB of flash program memory and 4kB RAM. Sending and receiving data in packet mode consume minimum of MCU resources. More problematic is time to transfer data via UART. Speed of ChipOx and ECG module is 9600 baud per sec. The communication speed of master with PC is 19200 baud per sec. Time to transfer data is too long and is necessary to use SCI receive and transmit interrupts and buffers not to oppress main program. Slaves are programmed to be expanded on new peripherals and master is programmed to easily add new slaves. All programming was done in Freescale Code Warrior in C language.

### 4 Other materials and methods

#### 4.1 Radio module

For data transmission from race car to depot were chosen OEM radio modules from Aerocomm company AC4868. These modules transmit at 868 MHz band with maximal output power 250mW, have very small dimensions (40 x 50 [mm]) and good maximal data rate (57.6 Kbps). These modules can work in acknowledged mode or stream data mode, data format is standard serial protocol. For this application was chosen stream data mode and maximal output power about 100mW, point – to – point operation mode. Inside race vehicle is small embedded system which provides communication between ZigBee module (master) and radio module. This system includes accelerometry measurement too, because it is fixated inside race vehicle. Transmitted data are received with used radio module, which is connected by serial interface to computer.

#### 5 Visualization

For data processing and visualization of measured values was created software called BioHydrogenix (see Fig. 4) in development system Labview 7.1. This software can process and visualized data from pulse oxymetry, ECG, temperature and used accelerometers. All measured data are saved to file, from which can be measured data easily imported to other software, like excel, for further analysis.
6 Results

It was designed and created wireless telemetry network of medical instruments, which are used for monitoring of base life function. The ZigBee technology was used for wireless data transmission. It shows that ZigBee can be used in much ambitious applications, like biotelemetry is. The power consumption of ZigBee slave is at about 30-35 mA while working. It can be lowered by increasing delays when MC13193 is suspended. When slaves are stopped they could have MC13193 suspended for longer period and search for masters’ signal less frequently to prolong lifetime.

This monitoring system was created for monitoring of pilot of race car with hydrogen power, and was tested this year at Shell Eco-marathon.

7 References

[5] Freescale: HC9S08GT60, MC13193 datasheet
USB real-time data acquisition module for ultrasound scanner

Martin Imramovský
VSB-Technical University of Ostrava, Czech Republic
martin.imramovsky@vsb.cz

Abstract: This paper introduces concept of specialized USB data acquisition module for real-time transfer of raw multi-modality data between ultrasound scanner and PC. Module allows its user to record, process and display raw images in the MATLAB and Simulink by special functions and blocks in real-time. Resulting images are not affected by additional noise and do not contain redundant information as it is common when capturing video signal.

1 Introduction

Research in the field of ultrasound signal and image processing requires availability of clinical data in the form of raw signals as well as images, echocardiograms and sonograms. There are several methods how the data can be obtained using commercially available ultrasound scanner, giving more or less suitable results for particular purposes (see [1] for more comprehensive discussion). This paper brings proposal of specialized USB add-on module for real-time transfer of raw multi-modality data between scanner and PC.

Using DICOM, mass storage (e.g. ZIP or CD-ROM) or video signal sampling (direct or with use of video tape recorder) has several disadvantages:

- original raw image samples cannot be obtained due to additional DSP (interpolation, averaging, grayscale modifications etc.)
- original Doppler signal cannot be obtained, thus spectrogram or CFM image cannot be calculated by different method
- during CFM, overlaying B-mode and flow images cannot be reconstructed
- ECG record cannot be obtained from images
- usually only static images or short loops can be recorded to scanner internal memory and then transferred to mass storage media or through DICOM
- all images incorporate useless supplementary data (they can even occupy one half of the image area)
- most scanners are not equipped with DICOM module

Scanner architecture conforms to the multimodality of such device. There are several types of data obtained by scanner:

- B-mode images
- D (Doppler) mode sonograms (i.e. spectrograms)
- CFM (Color Flow Mapping) images
- M-mode images (tissue motion)
- etc.

Scanner signal processing chain consists of many processing channels, i.e. B mode, M mode, D mode, CFM mode channel etc. Another channel introduces supplementary data into the
resulting image. They include patient record, examination notes, scales, various parts of user interface and ECG record (in the case of echocardiogram). Digital processing of image related signal includes interpolation, filtering, averaging, gray-scale coding and adjustment. D and CMF mode data are usually filtered, restricted to area of interest (i.e. sample volume), and processed with FFT for calculation of spectrogram (showing distribution of blood flow velocities). In the case of color flow mapping, calculated velocities are coded into color image. Using all types of processed data, image is then formed by overlaying them and converted into video signal or similar type of signal useful for displaying or recording. Each block of architecture can be realized by means of hardware and software, including embedded computer systems, FPGAs and digital signal processors. Detailed description of the ultrasound scanner signal processing chain can be found in [2].

2 Materials and methods

Image acquisition by envelope sampling is very useful when resulting images, echocardiograms and spectrograms are required. In the case of images, sampling rate is proportional to the number of obtained pixels and frame rate. For the highest frame rates (30 fps) used in basic clinical applications, sampling frequency will be maximally 10 MHz. When there is no need to capture all frames, it can be significantly lower. Sampling frequencies for M-mode and D-mode are in the order of kHz. The envelope is sampled on the output of receiver.

In the case of video signal sampling, resulting images are distorted due to extra digital-to-analog and analog-to-digital conversion by introducing quantization noise and due to induced noise. Another distortion is caused by unequal frame rate of ultrasound scanner and capturing device. Scanner frame rate vary with different ultrasound probes, due to changes in region of interest and imaging modes. Often, multimedia video capture card without trigger inputs are used, which makes capturing of correctly synchronized images impossible.

Video signal sampling and mass storage media are frequently used as an image source for research, e.g. in the area of image processing and recognition. The above mentioned disadvantages of such methods make them unusable as an source of valuable data. Unfortunately, they are often used in this manner. Raw images are more useful for image recognition methods than final interpolated ones, because of interpolation increases amount of data (as well as noise) and brings problems with triangular region of interest (if sector probe is utilized).

Our aim is to present less known approach to envelope sampling, which uses output of scanner internal ADC (or ADCs) and specialized USB module to transfer obtained raw data samples. Therefore, it provides uninterpolated images and raw Doppler signal. Fig. 1 shows basic block diagram of the proposed USB module. It was designed to meet several conditions:

- simultaneous transfer of multi-modality data
- real-time transfer and display of data in personal computer
- device should be able to set display modes and control other scanner settings
- maximal versatility of the device should allow to use it with several US models

Device is connected to ultrasound scanner using its internal control signals for transfer control. Important control signals slightly differ with scanner model, but have the same meaning:
Need of extremely fast microcontroller is eliminated by using programmable logic devices (PLDs) for low-level logic operations on control bus. Three main data buses (B/M, D/CFM, ECG) are alternately connected to FIFO through input multiplexer (MUX) owing to scanner control signals and device mode set by microcontroller. MUX is needed for consecutive transfer of diverse data types, which is realized during one transfer session. For example, echocardiogram is transferred as consecutive blocks of M-mode data and ECG data. This is permitted by relatively large break between consecutive image frames and lines (in all imaging modes). Therefore all three types of data can be successfully transferred during session. Often there are differences between bus sizes; B mode signal is sampled with 8, 10 or 12-bit ADC, however D mode signal is usually sampled with 16-bit ADC. Hence, input multiplexer must deal with this problem by reduction or expansion of bus size according to the size of the bus on the output of the device. The transfer is controlled by low-level logic. Another solution is to use FIFO with user-selectable input and output bus sizing (e.g. 16x8, 8x16). Through usage of FIFO memory, time breaks between consecutive image frames and lines are utilized to decrease required transfer speed of USB interface; in other words to fully utilize its transfer speed.

According to the type of USB interface, FIFO memory can be omitted if it is incorporated in such an interface. In most cases internal FIFO memories of the USB interfaces are very small, therefore stand-alone FIFO must be used. In order to simplify the whole design, FT245BM USB FIFO was employed in combination with another stand-alone FIFO. FT245BM (see [4] for more details) is single-chip USB to parallel interface with independent transmit/receive buffers and simple four-wire handshake interface. The entire USB protocol is handled internally and therefore this IC does not require any USB specific firmware. Drivers for most operating systems are provided for free. Application software can directly access device

Fig. 1 Basic block diagram of the add-on module.
buffers through USB by DLL based published API. Device is USB 2.0 compatible and its data transfer rate is up to 1 MBps.

It is advantageous to use FIFO memory with choice of memory organization (configurable input and output bus sizes). Independent read and write clock signals should permit simultaneous reading and writing. Our FIFO module consists of Texas Instruments SN74V2x3 FIFO, capable to store up to 128 kB. It can be easily expanded in both capacity and bus width by using more units [5].

Application software for the device was realized as a MATLAB toolbox and blockset for Simulink. Real-time displaying and registration of images is possible with use of specialized Simulink blocks, see Fig. 2. They serve as device output and viewers for basic imaging modes. Input block allows user to set scanner mode (type of transferred data) and USB transfer parameters, output FIFO timing and size of output frames. Simulink S-functions were created using C language and compiler. Source code employs DLL library provided by manufacturer of FT245BM to control USB interface and obtain its output. S-functions are compiled as C-MEX files, a kind of DLL libraries, utilizable by Simulink. Compilation has been done using internal MATLAB MEX utility and Microsoft Visual C/C++. Prototype of the proposed module was designed to be used with SonoAce 4800HD ultrasound scanner.

3 Results

The above proposed device was successfully realized and tested. Real-time imaging in B, M and A modes has been achieved using SonoAce scanner. It is possible to use both linear and convex probes. Fig. 3 shows examples of recorded B-mode images.

USB module was intended to be used in combination with analog beamformer scanner, but is easily adaptable to models with digital beamformer. Digital beamformers employ ADC in all parallel channels, digital adder and digital delay lines instead of analog ones. Device must be therefore connected behind the adder instead of ADC. However, it can be considerably harder to find appropriate point of connection.
Fig. 3 Images of liver obtained by USB data acquisition module were used as training set for neural network (from [3]).

4 Conclusions

This paper has shown possible methods how images acquired by commercially available ultrasound scanner can be obtained for further processing in PC. Disadvantages of video signal sampling and other methods led us to propose and realize specialized USB module for transfer of images and sonograms into PC. Device is very useful for testing a large family of possible image processing techniques. Need of raw unprocessed images, undistorted by noise, was a main motivation during device development. Device can also be helpful for better understanding of RF signal processing, providing raw images and data for comparison.

5 Acknowledgement

The support for this research work has been provided by the GAČR project 102/05/H525: The Postgraduate Study Rationalization at Faculty of Electrical Engineering and Computer science VSB-TU Ostrava.

6 References


Performance of Reduced Complexity Decoding of LDPC Codes with Few Quantization Bits

Wojciech Sulek
Silesian University of Technology
Wojciech.Sulek@polsl.pl

A reduced complexity iterative algorithm for decoding Low Density Parity Check (LDPC) codes is presented. The algorithm operates in the log-likelihood domain. Messages iteratively exchanged between symbol and parity-check nodes are log-likelihood ratios (LLR). In the practical hardware realization of a decoder, messages are represented by fixed-point numbers. For the purpose of area efficiency for the decoder, the number of bits per word shall be as low as possible. In this article there is shown, by simulations, that as few as 4-5 bits per message is sufficient for the decoder to operate correctly. The performance is reduced only a little in comparison with full SPA algorithm operating in the floating point double precision words.

1 Introduction

Low-Density Parity-Check codes are one of the two best known error correcting coding methods that are capable of achieving very low bit error rates at code rates approaching Shannon's channel capacity limit. (The second is turbo coding). LDPC codes were first introduced by Gallager in 1962 [1], but soon forgotten. Their implementation complexity was exceeding the capabilities of the accessible technology, so they were not considered for practical applications. The codes were rediscovered in the 90's [2] and after then they are under interests of many researchers. The reason is the need of the modern communication systems operating in high speed and low power, which means close to the Shannon limit.

LDPC codes are linear block codes defined by a sparse parity-check matrix \( H \) having \( M \) rows and \( N \) columns. Iterative decoding algorithm known as Two-Phase-Message-Passing (TPMP) algorithm can be described using a graph of the code constructed based on \( H \). Such a graph (Tanner graph - TG) is a bipartite graph - it contains two types of nodes, bit nodes and check nodes, corresponding to the columns and the rows of the matrix \( H \), respectively. A '1' located at particular position of \( H \) corresponds to an edge between respective bit-node and check-node. Iterative algorithms are performed by exchanging messages between bit-nodes and check-nodes through the edges in both directions. Each node corresponds to some mathematical operations. In the conventional algorithm, known as Sum-Product-Algorithm (SPA) [2], messages are extrinsic a posteriori probability values and operations are sum of products. This original algorithm is also considered to be the most complex iterative algorithm to implement and sensitive to quantization effects [3]-[4]. A number of simplifications of the SPA has been proposed [3]-[6] that reduces the high computational cost of the SPA at the price of some loss in performance.

In this paper a modified Min-Sum (MS) algorithm [3] with a double recursion scheme [4] is considered. The MS algorithm operates in the log-likelihood ratio (LLR) domain which means that logarithms of ratios of probabilities are used as messages. The algorithm offers a substantial reduction in complexity with essentially the same performance as the SPA. The paper is organized as follows. Firstly, LLR-SPA algorithm is described as well as its simplification – modified Min-Sum algorithm. Then double recursion scheme is presented,
which enables efficient implementation of the algorithm. At the end, simulation results of the fixed-point model of the algorithm are presented and some conclusions are drawn.

2 LLR-SPA decoding algorithm

If we denote the transmitted original codeword as $u=[u_1, u_2, \ldots, u_N]$ and the received corrupted by noise codeword as $y=[y_1, y_2, \ldots, y_N]$, then for each bit $u_n$ an *a posteriori* log-likelihood ratio (LLR) $L(u_n)$ is defined as:

$$
L(u_n) = \ln \frac{P(u_n = 0 | y_n)}{P(u_n = 1 | y_n)}
$$

Following the notation in [2]-[4] we denote the set of check nodes connected to symbol node $n$ by $M(n)$ and – similarly – $N(m)$ denotes the set of symbol nodes that participate in check $m$. Furthermore, $N(m)\setminus n$ represents the set $N(m)$ excluding the $n$-th symbol node and – similarly – $M(n)\setminus m$ represents the set $M(n)$ excluding the $m$-th check node. LLR messages from symbol node $n$ to check node $m$ are denoted as $\lambda_{n\rightarrow m}$ and messages from check node $m$ to symbol node $n$ are denoted as $\Lambda_{n\rightarrow m}$.

The LLR-SPA algorithm is then summarized as follows [4]:

**Initialization.** Each symbol node $n$ is assigned $L(u_n)$ value (in practical communication system these values are obtained from demodulator):

$$
\lambda_{n\rightarrow m} = L(u_n)
$$

$$
\Lambda_{m\rightarrow n} = 0
$$

**Step 1 (check-node update).** For each $m$, for each $n\in N(m)$, calculate messages $\Lambda_{m\rightarrow n}$ (separately sign and magnitude):

$$
|\Lambda_{m\rightarrow n}| = 2 \tanh^{-1} \left( \prod_{n\in N(m)\setminus n} \tanh \left( \frac{\lambda_{n\rightarrow m}}{2} \right) \right)
$$

$$
\text{sign}(\Lambda_{m\rightarrow n}) = \prod_{n\in N(m)\setminus n} \text{sign}(\lambda_{n\rightarrow m})
$$

**Step 2 (symbol node update).** For each $n$, for each $m\in M(n)$, calculate messages $\lambda_{n\rightarrow m}$:

$$
\lambda_{n\rightarrow m} = L(u_n) + \sum_{m'\in M(n)\setminus m} \Lambda_{m'\rightarrow n}
$$

**Step 3 (decision).** For each $n$, calculate:

$$
\hat{u}_n = L(u_n) + \sum_{m\in M(n)} \Lambda_{m\rightarrow n}
$$

then make hard decisions $\hat{u}=[\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_N]$ such that $\hat{u}_n = 0$ if $\lambda_n > 0$ otherwise $\hat{u}_n = 1$. If $\hat{u}H^T = 0$, then halt the algorithm with $\hat{u}$ as the decoder output, otherwise go to *Step 1*.

The most computationally complex operations are check-node updates which require tanh function calculation. In the next section the simplified version of check-node updates is described.
3 Modified Min-Sum algorithm with double recursion

3.1 Min-Sum (MS) algorithm

Several approximations for the core operation in check node processing (product of $tanh$ functions) have been published. One of the most computationally effective algorithm is a “Min-Sum” (MS) algorithm [3].

The magnitude of messages in this algorithm is calculated as follows. For each $m$, for each $n \in N(m)$:

$$|\Lambda_{m\rightarrow n}| = \min_{n \in N(m)} \left( |\lambda_{n\rightarrow m}| \right) \quad (6)$$

The other steps and initialization are exactly the same as in SPA – see (2)-(5).

3.2 Modified Min-Sum algorithm

MS algorithm has low complexity at a cost of reduced performance – it requires a few tenths of a dB more transmitted power than SPA ([3]). Some modifications can be made to improve the performance with small increase in complexity. These modifications have a form of a correction factors that are typically added (subtracted) to the result of (6).

Suppose we have a check node $m$ with degree 3, connected with symbol nodes $n_1$, $n_2$, $n_3$. Then the output message formula (3) can be written in other way as:

$$|\Lambda_{m\rightarrow n}| = \frac{1+e^{B_{2\rightarrow m}}e^{B_{3\rightarrow m}}}{e^{B_{2\rightarrow m}} + e^{B_{3\rightarrow m}}} \quad (7)$$

Similarly the other two output messages can be calculated.

Now, assuming that $e^{B_{2\rightarrow m}}e^{B_{3\rightarrow m}} = 1$, above equation can be simplified as:

$$|\Lambda_{m\rightarrow n}| = \min \left( |\lambda_{n_2\rightarrow m}|, |\lambda_{n_3\rightarrow m}| \right) + \ln \left( \frac{e^d}{e^d + 1} \right) = \min \left( |\lambda_{n_2\rightarrow m}|, |\lambda_{n_3\rightarrow m}| \right) + f(d) \quad (8)$$

$$d = |\lambda_{n_2\rightarrow m}| - |\lambda_{n_3\rightarrow m}|$$

This is the form that was mentioned above – a Min operator with correction factor $f(d)$ which is a nonlinear function of difference of input messages.

3.3 Double recursion scheme

To extend the above algorithm for check nodes with degree higher than 3, a double recursion scheme can be proposed [4]. We define two auxiliary sets of variables: $\alpha_i$ for “forward” recursion and $\beta_i$ for “backward” recursion. The modified Min-Sum algorithm with double recursion scheme can then be given as follows:

$$\alpha_{i+1} = \min \left( \alpha_i, |\lambda_{n_i\rightarrow m}| \right) + f(d), \quad i = 1 \ldots (d_c - 2)$$

$$\beta_i = \min \left( \beta_{i+1}, |\lambda_{n_i\rightarrow m}| \right) + f(d), \quad i = (d_c - 3) \ldots 0$$

$$|\Lambda_{m\rightarrow n}| = \min (\alpha_i, \beta_i), \quad i = 0 \ldots (d_c - 1) \quad (9)$$

where variables are initialized as: $\alpha_0 = |\lambda_{n_1\rightarrow m}|$, $\alpha_1 = |\lambda_{n_2\rightarrow m}|$, $\beta_{d_c - 1} = |\lambda_{n_2\rightarrow m}|$, $\beta_{d_c - 2} = |\lambda_{n_1\rightarrow m}|$; $d_c$ denotes the check-node degree; correction factor $f(d)$ is as defined in (8).
Gain of using the recursion scheme is twofold. First, it enables incorporating a correction as in (8) for Min-Sum algorithm. Second, it requires fewer operations in comparison with “brute-force” implementation of (6). Direct implementation of (6) – for each \( n \in N(m) \) separately – possesses some redundant operations that are eliminated when using recursion scheme (9).

Block diagram of the forward recursion hardware unit is shown in Figure 1. (Similarly backward recursion unit and final-pairwise (9) can be realized). Subtractor on the left of the figure is used for Min operator (together with Multiplexer) as well as for addressing the Look-Up-Table (LUT). The LUT outputs the value of correction factor \( f(d) \).

![Block diagram of forward recursion unit](image)

**4 Fixed-point decoding – simulation results**

In this paragraph, the finite precision analysis of the algorithm is presented. In a practical hardware realization of a decoder, messages are represented by fixed-point numbers. For the purpose of area efficiency for the decoder, the number of bits per word shall be as low as possible. On the other hand, the fewer bits per message – the lower performance of the algorithm.

In order to analyze the effects of message quantization on the performance, a simulation of the decoding algorithm with finite precision fixed point numbers shall be done. The simulations were made using the Matlab environment. The fixed-point model that have been created corresponds exactly to the modified Min-Sum algorithm with double recursion scheme (equation (2) for initialization, eq. (9) and Fig. 1 for check-node update, eq. (4) for symbol node update and eq. (5) for decision rule). All messages are represented by fixed-point numbers, with two parameters: number of bits \( q \) and clipping threshold \( c_{th} \). In the corresponding hardware implementation, convenient way of representing the numbers is “sign and magnitude” form. It means that there are \( 2^q - 2 \) quantization levels uniformly placed in the range \([-c_{th}, c_{th}]\). The signal granularity is thus given by: \( \Delta = 2c_{th}/(2^q - 2) \).

A number of simulations were made in order to draw some conclusions about the required number of bits \( q \) and clipping threshold \( c_{th} \). Figures 2-3 contain results for regular (1000, 500)
LDPC code with no length-4 cycles. Simulations were made for binary NRZ modulation and AWGN channel model. Maximum number of iterations was restricted to 80.

Figure 2 contains Bit Error Rate (BER) curves for LLR-SPA as well as for our fixed point model of modified Min-Sum algorithm with $q = 4$, $q = 5$, $q = 7$ and clipping with $c_{th} = 12$. When messages are quantized with more than 7 bits, improvement in performance is very little. When fewer than 5 bits are used, the loss in performance is quite significant. Thus reasonable values of $q$ are between 5 and 7.

Figure 3 presents influence of clipping threshold $c_{th}$ on the performance. It shows that this influence should not be neglected. The best results (for $q = 5$) were obtained with $c_{th} = 6$. This combination of parameters ($q = 5$, $c_{th} = 6$) can be considered as good compromise between performance and complexity of the decoder. The loss in performance is quite little in this case.

![BER curves for (1000,500) regular LDPC code decoded by LLR-SPA as well as modified Min-Sum algorithm with quantization and clipping with $c_{th} = 12$](image)

**5 Conclusions**

The modified Min-Sum algorithm with double recursion scheme for decoding LDPC codes was presented in the paper. The algorithm is convenient to implement in hardware like Programmable Logic Devices.

Careful analysis of the messages quantization effects is a must when practical hardware realizations are developed. Simulation results show that a 5-bit quantization with precise selection of the clipping threshold appears to be a good compromise between complexity and performance. The optimum clipping threshold depends on the code. Another simulations show that for some codes as low as 4-bit quantization is sufficient. Thus, a general conclusion can be made that depending on the code, 4–5 bits per message are required and clipping threshold shall be in the range of 5…10.
Fig. 3  BER curves for (1000,500) regular LDPC code decoded by modified Min-Sum algorithm with 5-bit word and clipping with different threshold $c_{th}$

6 References