

REZUMATUL TEZEI DE ABILITARE

**STUDII ȘI CERCETĂRI PRIVIND
PRELUCRAREA MODERNĂ A SEMNALELOR
BIOMEDICALE DE TIP ECG ȘI EEG**

Domeniul de abilitare: Electronica, Telecomunicații și Nanotehnologie

Autor: FIRA CATALINA MONICA

STUDIES AND RESEARCH ON THE MODERN PROCESSING OF ECG AND EEG BIOMEDICAL SIGNALS

This habilitation thesis presents the main results obtained and published by the candidate after obtaining the PhD degree in 2006. The doctoral thesis was entitled "*Contributions to biomedical signal processing*" and was conducted between 2002 and 2006 under the supervision of Prof. Dr. Eng. Liviu Goraş. The research activities following to the doctoral dissertation were carried out in the field of signal processing by mean of modern techniques of biomedical signals of ECG and EEG type, with applicability in two main fields, namely, compression and classification of signals.

In the biomedical signal compression field, two main directions are identified, which were addressed by the candidate in his research activities and are presented in five chapters of this thesis, namely: compression of ECG signals, compressed sensing of ECG signals using dictionaries specific to these signals, compressed sensing of EEG signals using dictionaries specific to these signals and reduction of the signal dimensionality with new techniques such as Locality Preserving Projections - LPP and Laplacian Eigenmaps - LE.

In the signal classification field, two major directions are identified, namely, the classification of normally acquired signals and the classification of biomedical signals that have a smaller dimensionality than the initial one, i.e., compared to that imposed by the sampling theorem.

The most important results obtained by the candidate in the research activities were validated by publication in several prestigious international journals listed Q1 and Q2 in the WOS database, (IEEE Transactions on Biomedical Engineering and Biosensors-Basel), as well as other journals with impact factor or in volumes of prestigious international conferences in the research fields of interest.

Chapter 1 presents some of the author's research on how to preprocess ECG signals in order to compress and classify these signals. Some techniques for ECG signals preprocessing and segmenting are presented, as well as the influence of features and the database used in classification. The topics presented in this chapter can be outlined in three different directions, namely, (i) presenting a new solution for ECG signal compression, (ii) improving the classification of multiclass data, and (iii) features selection based on the Laplacian score.

Chapter 2 presents the author's research in the field of compressed acquisition with applicability on electrocardiographic (ECG) biomedical signals. Thus, a detailed comparative study of two different approaches for the compressed acquisition of the ECG signals is presented. This study considers several acquisition techniques and several types of projection metrics used in the acquisition phase, as well as several dictionaries used in the ECG signal reconstruction phase. The effect of preprocessing on the results is also analysed. For both variants proposed by the author, the influence of the projection matrix in the final results was analysed, namely, several projection matrices were analysed: matrix with independent random elements and identically distributed (i.i.d.) with Gaussian distribution, Bernoulli type matrix, and projection matrix optimized for the private dictionary used in reconstruction. In addition, special attention was given to the way the dictionary is constructed, both methods having different ways of constructing. The advantages and disadvantages of each method and the choice of a method according to the available hardware and software resources are also presented.

Chapter 3 focuses on how to build dictionaries used in the reconstruction of the compressed acquired EEG signals. Similar with the ECG signals, dictionaries can be built for EEG signals that take into account the statistics and features of this type of signal or the patient's particularities. Starting

from the alignment idea, which in the case of the ECG signal led to much improved results, for the EEG signal an alignment method was analysed and tested. Various dictionaries with EEG segments have been built, such as the patient-specific dictionary, the channel-specific dictionary, a predefined number of temporal specific dictionaries, or the Daubechies10 wavelet dictionary. For all these dictionaries, the obtained reconstruction results were compared. Another analysis was the influence of the acquisition matrix used in the reconstruction results. The conclusions of the analysis were that for small compressions (e.g. CR = 5: 1) the dictionary-based method with aligned EEG segments according to the stimulus in the spelling paradigm provided the best results. However, for large compressions (above 10: 1) the best method of acquisition is based on the dictionary with atoms of 1 sec size with P300 wave and NonP300 atoms. The disadvantage of this method is that the method also requires a realignment step for the reconstructed signals. Another definite conclusion is that the worst results are obtained when standard wavelet dictionaries were used. In the case of the universal mega-dictionary and the intra- and inter-patient analysis, the best results were obtained when the dictionary and test data came from the same subject but different EEG signals were used (i.e. for the dictionary EEG signals from the training phase were used and for the compressed acquisition EEG signals from the testing phase were used).

Chapter 4 presents the classification performance for electrocardiographic (ECG) and electroencephalographic (EEG) signal classes processed for different degrees of dimensionality reduction. The results obtained with different classification methods were analysed and discussed. Three dimensionality reduction techniques were investigated, namely: Laplacian Eigenmaps (LE), Locality Preserving Projections (LPP) and compressed sensed (CS). The first two methods are related to multiple learning, while the third refers to the acquisition and reconstruction of the signal from random projections under the assumption of signal sparsity. The aim of the study was to evaluate the benefits and disadvantages of the different methods and to find out to what extent they can be considered remarkable. The evaluation of the effect of dimensionality decreasing was made taking into account the classification rates for the biosignals processed in the new spaces. In addition, the accuracy of the classification of the initial input data was assessed against the corresponding accuracy in the new spaces using different classifiers.

The main conclusions of the analysis regard the choice of the combination of dimensional reduction techniques and classification algorithms to achieve reasonable classification results even for small or very small dimensions for both ECG and two classes of EEG signals. The choice of the reduction rate depends on the motivation of the analysis. Thus, if it is wanted to reconstruct the initial signal, the CS method will be adopted; if the reduction of the space to 2D or 3D is targeted in order to understand the data, the LE method will be chosen, while if the goal is to reduce the dimensionality by about 10 to 12 times or to classify the new data in the reduced space without recalculating the initial signals, the LPP method will be used.

Chapter 5 presents the possibility of classifying biomedical signals purchased in tablets directly in this small space, without the need to reconstruct the ECG or EEG signal. The classification in compressed space for the ECG signal is an important advantage when using dictionaries specific to the pathologies for the reconstruction of the original signals, because once identified the class to which the pattern belongs allows the use of that dictionary specific to that class. Another important aspect is that with the acquisition of a compressed signal based on a random matrix, the classification of the compressed signal provides additional information about the nature and type of the original signal. Thus, both compressed and original cardiac models were classified. Similar results were obtained in the original space with those in the compressed space. To validate the correctness of the

results, two classifiers were tested, an MLP and a KNN. Thus, it is found that if an advantageous architecture is chosen for the MLP network, it is found that the two classifiers behave similarly in terms of classification performance.

For EEG signals and acquisition in compressed space topic, the possibility of EEG signals classifying (from the spelling paradigm) into EEG segments containing P300 waveform and EEG segments without P300 waveform was analysed, the classification taking place directly on the compressed acquired segments. This classification is the key element in a BCI spelling system. Thus, starting from a method proposed by Hoffmann, based on gradient boosting, the possibility of classifying the collected EEG signals was tested. The possibility of classification using Deep Learning neural networks was also studied, the results obtained in terms of classification being very close to those got using the gradient boosting approach. The results obtained with both tested methods confirm the hypothesis that close neighbourhoods in the initial space remain close also in the compressed space. This result allows the classification of the directly compressed acquired signals and it is useful in medical applications where only the class of a signal is necessary and not its shape.

Section II of the thesis illustrates the evolution and development of the candidate's academic career from the perspective of research activity. This chapter highlights the most important research directions that will be addressed in the field of signal processing and how to attract research funding by participating at competitions either in national and international research or for improving the equipment of laboratories.

The candidate's experience gained during her scientific research career as well as the experience gained as a project director or member of the implementation team or through numerous scientific papers published in prestigious journals and international conferences, will focus on strengthening the research team by attracting new members, especially from doctoral students. Thus, the attraction of students will be achieved through the good collaboration of the candidate with university professors from the Faculty of Bioengineering and the Faculty of Electronics, Telecommunications and Information Technology. This fruitful collaboration is proved by the scientific papers published together with professors from those faculties but also by the collaboration of a PN II Partnerships project - coordinated by UMF Iași in partnership with the Romanian Academy - Theoretical Informatics Institute and Gh. Asachi Polytechnic University - Faculty of Electronics, Telecommunications and Information Technology.

*
* *
* * *

I thank Prof. Dr. Eng. Liviu Goraș for the support and advice given during the elaboration of the doctoral thesis, the habilitation thesis, as well as for the constant support throughout this period. He was like a third parent through the advice and patience he gave me in our research activities done together.

I thank Acad. Horia-Nicolai Teodorescu and Director Dr. Vasile Apopei for the good collaboration and development of the research activity within the Institute of Theoretical Informatics.

I would like to thank all the collaborators over the years for their support in the research grants and for their involvement in many scientific papers developed and published by our team.

I am also grateful to the entire staff of the Institute of Theoretical Informatics for facilitating a pleasant atmosphere in terms of research activity.

Special thanks to my family for their support and understanding over the years.

REZUMATUL TEZEI DE ABILITARE

**STUDII ȘI CERCETĂRI PRIVIND
PRELUCRAREA MODERNĂ A SEMNALELOR
BIOMEDICALE DE TIP ECG ȘI EEG**

Domeniul de abilitare: Electronica, Telecomunicatii și Nanotehnologie

Autor: FIRA CĂTĂLINA MONICA

Candidat: CSII dr. bioing. Cătălina Monica Fira

Teza de abilitare STUDII ȘI CERCETĂRI PRIVIND PRELUCRAREA MODERNĂ A SEMNALELOR BIOMEDICALE DE TIP ECG ȘI EEG

Teza de abilitare prezintă principalele rezultate obținute și publicate după obținerea titlului de doctor inginer în anul 2006 cu teza de doctorat "*Contribuții la prelucrarea semnalelor biomedicale*", realizată în perioada 2002 - 2006 sub îndrumarea Prof. Dr. Ing. Liviu Goraș. Activitățile de cercetare ulterioare susținerii doctoratului au vizat domeniul prelucrării cu ajutorul tehnicilor moderne a semnalelor biomedicale de tip ECG și EEG, cu aplicabilitate în două domenii principale, și anume, compresia și clasificarea semnalelor.

În domeniul compresiei semnalelor biomedicale am abordat două direcții principale, care sunt prezentate în cinci capitole ale tezei, și anume: compresia semnalelor ECG; achiziția comprimată a semnalelor ECG cu ajutorul dicționarelor specifice acestor semnale; achiziția comprimată a semnalelor EEG cu ajutorul dicționarelor specifice acestor semnale; reducerea dimensionalității semnalelor cu tehnici noi precum mape proprii Laplaciene (Laplacian Eigenmaps – LE), proiecții de conservare a localizării (Locality Preserving Projections - LPP).

În domeniul clasificării semnalelor am abordat: clasificarea semnalelor achiziționate normal; clasificarea semnalelor biomedicale care au o dimensionalitate redusă față de cea inițială, adică față de cea impusă de teorema eșantionării.

Cele mai importante rezultate obținute în activitățile de cercetare au fost validate prin publicare în câteva reviste internaționale de prestigiu cotate Q1 sau Q2 în baza de date WOS, (IEEE Transactions on Biomedical Engineering și Biosensors-Basel), precum și alte reviste cu factor de impact sau în volumele unor conferințe internaționale de prestigiu din domeniile de cercetare de interes.

În **Capitolul 1** am prezentat câteva cercetări proprii privitor la modalitatea de preprocesare a semnalelor ECG în vederea compresiei și clasificării acestor semnale. Am prezentat câteva tehnici de preprocesare și segmentare a semnalelor ECG, precum și influența trăsăturilor și a bazei de date folosită în clasificare. Subiectele prezentate în acest capitol se pot schematiza pe următoarele direcții: (i) prezentarea unei noi soluții pentru compresia semnalului ECG; (ii) îmbunătățirea clasificării datelor multiclase; (iii) selecția trăsăturilor bazată pe scorul Laplacian.

În **Capitolul 2** am prezentat cercetările proprii în direcția achiziției comprimate cu aplicabilitate pe semnalele biomedicale de tip electrocardiografic (ECG). Astfel, am prezentat un studiu comparativ detaliat a două abordări proprii diferite de achiziție comprimată a semnalelor ECG. Acest studiu ia în considerare mai multe tehnici de achiziție și mai multe tipuri de metrici de proiecție utilizate în etapa de achiziție, precum și câteva dicționare utilizate în etapa de reconstrucție a semnalelor ECG. De asemenea, am analizat efectul preprocesării asupra rezultatelor. Pentru ambele variante propuse am analizat și influența matricei de proiecție în rezultatele finale, și anume, am analizat mai multe matrici de proiecție: matrice cu elemente aleatoare independente și distribuite identic (i.i.d.) cu distribuția Gaussiană; matrice de tip Bernoulli și matrice de proiecție optimizate pentru dicționarul particular utilizat în reconstrucție. Am acordat o atenție deosebită modului în care este construit dicționarul, prin prezentarea a două metode de realizare, ambele metode având modalități diferite de construire. De asemenea, am prezentat avantajele și dezavantajele fiecărei metode și alegerea unei metode funcție de resursele hardware și software disponibile.

Capitolul 3 se concentrează pe problematica modalității de construire a dicționarelor utilizate în reconstrucția semnalelor EEG achiziționate comprimate. La fel ca și în cazul semnalelor ECG, pentru semnalele EEG se pot construi dicționare care să țină cont de statistica și particularitățile

acestui tip de semnal sau particularitățile pacientului. Plecând de la idea de aliniere, care în cazul semnalului ECG a condus la rezultate mult îmbunătățite, pentru semnalul EEG am analizat și testat o modalitate de aliniere. Am construit diferite dicționare cu segmente EEG precum, dicționar specific pacientului, dicționar specific canalului, un număr prestabilit de dicționare specifice temporale sau dicționar *wavelet Daubechies10*. Pentru toate aceste dicționare am comparat rezultatele de reconstrucție obținute. O altă analiză a fost influența matricei de achiziție utilizate în rezultatele de reconstrucție. Concluziile analizei au fost că „pentru compresii mici (de ex de CR = 5:1) este mai bună metodă bazată pe dicționar cu segmente EEG aliniată funcție de stimulul din paradigma de spelling. Pentru compresii mari (de peste 10:1) este mai bună metoda de achiziție comprimată bazată pe dicționar cu atomi de dimensiune 1 sec cu unda P300 și atomi NonP300”. Dezavantajul acestei metode este însă faptul că necesită și o etapă de realiniere a semnalelor reconstruite. O altă concluzie certă este că „cele mai slabe rezultate se obțin atunci când s-au folosit dicționare wavelet standard”. În cazul mega-dicționarului universal și a analizei intra- și inter-pacient, cele mai bune rezultate s-au obținut când dicționarul și datele de test au provenit de la același subiect însă s-au folosit semnale EEG diferite (adică pentru dicționar s-au folosit semnale EEG din faza de antrenare și pentru achiziție comprimată s-au folosit semnale EEG din faza de testare).

Capitolul 4 prezintă performanțele de clasificare pentru clase de semnale electrocardiografice (ECG) și electroencefalografice (EEG) procesate pentru diferite grade de reducere a dimensionalității. Rezultatele obținute cu diferite metode de clasificare au fost analizate și discutate. Am investigat trei tehnici de reducere a dimensionalității, și anume: mape proprii Laplaciene (Laplacian Eigenmaps - LE), proiecții de conservare a localizării (Locality Preserving Projections - LPP) și achiziție comprimată (compressed sensed - CS). Primele două metode sunt legate de învățarea multiplă, în timp ce a treia se referă la achiziționarea și reconstrucția semnalului din proiecții aleatorii sub presupunerea rarefierii semnalului. Scopul studiului a fost de a evalua beneficiile și dezavantajele diferitelor metode și de a afla în ce măsură acestea pot fi considerate remarcabile. Evaluarea efectului scăderii dimensionalității a fost făcută luând în considerare ratele de clasificare pentru biosemnalele procesate în noile spații. În plus, acuratețea clasificării datelor de intrare inițiale a fost evaluată în raport cu acuratețea corespunzătoare în spațiile noi folosind diferiți clasificatori.

Principalele concluzii ale analizei vizează modul în care algoritmi de reducere a dimensionalității și de clasificare pot fi combinați pentru a obține rezultate rezonabile de clasificare chiar și pentru dimensiuni mici sau foarte mici atât pentru ECG cât și pentru două clase de semnale EEG. Alegerea ratei de reducere a dimensionalității depinde de scopul analizei. Astfel, dacă se dorește reconstruirea semnalului inițial, se va adopta metoda CS; dacă se dorește reducerea spațiului la 2D sau 3D cu scopul înțelegerii datelor, se va alege metoda LE, în timp ce dacă se dorește să se reducă dimensionalitatea de aproximativ zece - douăsprezece ori și să se clasifice noile date în spațiul redus fără a recalcula semnalele inițiale, se va folosi metoda LPP.

Capitolul 5 prezintă posibilitatea clasificării semnalelor biomedicale achiziționate comprimat direct în acest spațiu redus, fără a fi nevoie să se reconstruiască semnalul ECG sau EEG. Clasificarea în spațiu comprimat pentru semnalul ECG este un avantaj important în cazul utilizării de dicționare specifice patologiilor pentru reconstrucția semnalelor originale, deoarece odată identificată clasa din care face parte *patternul* permite utilizarea aceluși dicționar specific clasei respective. Un alt aspect important este că odată cu achiziționarea comprimată a unui semnal pe baza unei matrice *random*, prin clasificarea semnalului comprimat se cunosc informații suplimentare despre natura și tipul semnalului original. Astfel, am clasificat atât modelele cardiace comprimate cât și originalele. Am obținut rezultate similare în spațiu original cu cele din spațiu comprimat. Pentru validarea

corectitudinii rezultatelor, am testat doi clasificatori, un MLP și un KNN. Am constatat că dacă se alege o arhitectură avantajoasă pentru rețeaua de tip MLP se constată că cei doi clasificatori se comportă asemănător ca și performanțe de clasificare.

Pentru semnale de tip EEG și achiziția în spațiul comprimat, am analizat posibilitatea clasificării semnalelor EEG (provenite din cadrul paradigmei de spelling) în segmente EEG care conțin forma de unda P300 și segmente EEG fără unda P300, clasificarea având loc direct pe segmentele achiziționate comprimat. Aceasta clasificare este elementul esențial în cadrul unui sistem *BCI de spelling*. Astfel, plecând de la o metoda propusa de Hoffmann, bazată pe *gradient boosting* am testat posibilitatea clasificării semnalelor EEG culese comprimat. Am studiat și posibilitatea clasificării utilizând rețele neuronale de tip *Deep Learning*, rezultatele obținute din punct de vedere al clasificării fiind foarte apropiate de cele obținute cu gradient boosting. Rezultatele obținute cu ambele metode testate confirmă ipoteza conform căreia vecinătăți apropiate în spațial inițial rămân apropiate și în spațial comprimat. Aceasta permite clasificarea semnalelor achiziționate direct comprimat și este utilă în aplicații medicale unde importantă este doar clasa de apartenență a unui semnal și nu forma acestuia.

Sectiunea II a tezei ilustrează evoluția și dezvoltarea carierei academice din perspectiva activității de cercetare. În acest capitol am evidențiat cele mai importante direcții de cercetare pe care le voi aborda în domeniul procesării semnalelor și modalitatea de a atrage fonduri de cercetare prin participarea la competiții naționale și internaționale de cercetare sau cu caracter de îmbunătățire a dotărilor laboratoarelor.

Experiența câștigată de-a lungul carierei științifice de cercetare precum și experiența câștigată ca director de proiect sau membru în echipa de implementare, însă și cea acumulată prin numeroasele lucrări științifice publicate în reviste și la conferințe internaționale prestigioase, va fi concentrată pe consolidarea echipei de cercetare prin atragerea de noi membri, în special din rândul studenților la doctorat. Astfel, pentru atragerea studenților voi îmbunătăți colaborarea, deja existentă, cu profesori universitari de la Facultatea de Bioinginerie a Universității de Medicina și Farmacie „Gr. T Popa” (UMF) și Facultatea de Electronică, Telecomunicații și Tehnologia Informației a Universității de Tehnice „Gheorghe Asachi” din Iași (UTI). Această colaborare fructuoasă este dovedită și prin lucrările științifice publicate împreună cu profesori de la aceste facultăți dar și prin colaborarea la un proiect PN II Parteneriate - coordonat de UMF Iași în parteneriat cu Academia Romana - Institutul de Informatică Teoretică și UTI.

*
* *
* * *

Mulțumesc domnului prof. dr. ing. Liviu Goraș pentru suportul constant și sfaturile acordate în timpul elaborării tezei de doctorat, a tezei de abilitare, precum și în toată perioada de colaborare. A fost ca un al treilea părinte prin sfaturile și răbdarea pe care mi le-a acordat în activitatea de cercetare desfășurată împreună.

Mulțumesc domnului Acad. Horia-Nicolai Teodorescu și domnului director Vasile Apopei pentru buna colaborare și desfășurare a activității de cercetare din cadrul Institutului de Informatică Teoretică.

Doresc să adresez mulțumiri tuturor colaboratorilor de-a lungul anilor pentru suportul acestora în cadrul granturilor de cercetare și pentru implicarea în numeroase lucrări științifice elaborate și publicate de echipa noastră.

Sunt de asemenea recunoscătoare întregului colectiv al Institutului de Informatică Teoretică pentru facilitarea unei atmosfere plăcute în ceea ce privește activitatea de cercetare.

Adresez mulțumiri speciale familiei mele pentru suport și înțelegere de-a lungul anilor.