

# COMPRESSION AND INTERPOLATION PROBLEMS FOR ECG MAPPING

**T. L. Dragomir, L. Stoicu-Tivadar, S. A. Frantescu**

Department of Automation and Industrial Informatics

Faculty of Automation and Computers

"Politehnica" University of Timi<sup>o</sup>oara, Romania

*Abstract: The paper presents several problems of compression and interpolation associated with measurements in mapping ECG. There are presented two alternate methods of compression and interpolation and their implementation in processing environments. A comparative study between the results yielded applying the two methods will be done.*

*Keywords: compression, interpolation, mapping, RIP, significant points, comparison.*

## 1 INTRODUCTION

Given the current situation in health problems, where the coronary diseases are on the first place, the multitude of studies connected with the domain is somehow understandable. Standard 12 leads ECG and ECG mapping are two instrumentation techniques that yield data that, after processed in characteristic modes, give representations useful in heart diagnosis. The processing in ECG mapping are mainly based on compression and interpolation methods.

In this context the paper presents aspects referring two procedures of data compression and interpolation. The first is based on the procedure presented in detail in [1]. This completes and develops an idea presented in [2]. The second procedure is based on the RIP method [3] that in the step of rough interpolation uses the method of interpolation Shepard [4]. In the second case, we deal with the extrapolation of a technique from the domain of control to the domain of mapping ECG, as a mean of compression and reconstruction of signals. For both procedures there are used specific developing environments. These were conceived either dedicated to the problem of mapping ECG [2] either for the general case [5].

The contributions of the paper are at the conceptual and also application level and the results are at least of interest. The presentation of the subject is organised in four parts. The frame of the paper and the problems of principle referring the ECG mapping are presented in section 2. The first method, based on considering the measured potentials as a stochastic process is exposed in section 3. The second method, based on Shepard interpolation is developed in section 4. Finally, in section 5 is made an evaluation of the methods based on the results yielded using the development environments created by the authors.

## 2 MEASUREMENTS AND PROCESSING PROBLEMS IN ECG MAPPING

The technique of ECG mapping is an alternative for the standard 12-lead ECG technique. As a particular case in mapping, the *body surface potential maps* (BSPM) representation and processing is part of a very complex process which can constitute a shortcoming in using them. Another, more general representation, the *potential surfaces from the torso* (PST), imply the same problems.

### 2.1. The utility of an algorithm for the significant points

The complex built with all the ECG registrated from a grid on the human torso (all the BSPM), for different patients is a stochastic process (on the surface and in time) also for one subject or for a group of subjects. This is sustained by the time and surface variation of the signals registrated on the surface (for an individual) and by the different torso shapes for different individuals (from a group).

The data amount resulted from this type of investigations is great. As an *example*: for 112 grid nodes and 700 samples for a cardiac cycle we will have  $112 \times 700 = 78\ 400$  values/cardiac cycle. For every moment of time a BSPM can be drawn and we will have 700 BSPM. For every node an ECG can be visualised and we will have 112 ECG signals. Studying these various and numerous representation and establishing a correct diagnosis based on them is extremely inconvenient and difficult. From practical reasons in connection with the data storage, process and interpretation, hardware and instrumentation

problems, new computing and interpretation methods, which can drive to correct diagnosis, must be developed. An important solution for these problems is the methods for data compression and rebuild.

In order to have the advantages of working with BSPM/PST and also the ones of an acquisition and instrumentation system with a reduced number of electrodes there were developed methods through which are selected certain measure points in a grid that constitutes the base for the BSPM and from the information yielded by this points will be computed the values in the estimation points (the points from the grid where no registration are done). The first step means a data reduction and the second a data extension.

The algorithm that yields the significant points (SP) represents a non-standard approach for the domain of electrocardiography (a stochastic approach) which results in easy to use representations, very proper for a physician for the diagnosis process (even that the way of obtaining the results has a high rate of difficulty). The compression can be seen on surface, when we make processing only for a moment from the cardiac cycle, or surface and temporal compression/extension when the processing is made for all the leads in all the moments of the cardiac cycle.

For overcoming the disadvantage of the complexity of acquisition and processing of the information contained by BSPM, an algorithm to obtain the global optimal set of significant points (GOSSP) it is suggested using a programming environment that yields several sets of significant points (SP) and, based on the minimal error, the GOSSP. Using the GOSSP has two important effects:

- reduces the *complexity* of the acquisition and instrumentation system;
- smaller instrumentation system means an *enhanced comfort* for the patient during the investigations.

The paper presents results yielded by two methods:

- one with a more accurate mathematical support, based on stochastic analysis that drove to the GOSSP (§ 3.1);
- one based on RIP technique (§ 4).

In both cases were used specific development environments [2], [5] that are based on algorithms with a common macrostructure, presented in 2.2.

## 2.2 The steps of the algorithm based on stochastic characteristics associated to specific development environments

Connected to the SP two major problems are emphasised:

- how to select **the number and the positions** of the SP (p);
- which is the **transformation** through which the values in the estimation points (EP) are obtained from the values in the SP (S).

The process to solve these problems implies two steps:

- (a) the **development step** in which, starting from a great number of measurements are established the number and positions of the SP and the transformation;
- (b) the **application step** - the results from the first step are used in current practice: from the measurements in the SP the values in the EP are estimated and then the BSPM are drawn for a cardiac cycle and based on that the health state of the heart can be evaluated.

The method that uses stochastic representations implies processing of a great amount of data resulted from ECG acquisition in the nodes of the grid drawn on the human torso. Its advantages are new processing and representation possibilities useful in diagnosis of the state of the heart without the computing effort previously mentioned and with the same error level.

## 3 STATISTICAL PROCESSING AND OPTIMIZATION IN ECG MAPPING COMPRESSION AND EXTENSION. CASE STUDY

### 3.1 The global optimal set algorithm

A technique is suggested [2] through which from a lot of sets of SP [1] is chosen the **optimal global set**. This technique is recommended for computers with high computing speed rates and can process great amounts of data. Considering a set of  $n$  measuring points for BSPM registration, the **significant points** are the set of  $n_{PS} \leq n$  points that contain the essential information about the health state of the heart. The number of SP will remain fixed. Once the SP set established, measurements in the SP and a transformation give the values in the rest of the  $n - n_{PS} = n_{PE}$  points, called estimation points and offer the possibility to redraw the BSPM.

The algorithm for SP set selection finds an optimal subset of SP that minimizes the estimation error of the process  $I_e$

$$I_e = \text{tr} \left( E \left[ \left( P_{PE} - \hat{P}_{PE} \right) \left( P_{PE} - \hat{P}_{PE} \right)^T \right] \right) \quad (1)$$

$P_{PE}$  is the vector formed with the potential registered in the EP and  $\hat{P}_{PE}$  is the estimation for  $P_{PE}$ . After computations results [2]:

$$I_e = \text{tr}(C) - \text{tr}(C_{11}^{-1} L_{11}) \quad (2)$$

where

$$L_{11} = C_{11} C_{11} + C_{12} C_{12}^T \quad (3)$$

and  $C$  is the covariance matrix of the process

$C_{ij}$  its partitions from the SP and the EP.

As  $\text{tr}(C)$  has a constant value in order to minimize  $I_e$  is necessary to maximize the second term of (2),  $\text{tr}(C_{11}^{-1} L_{11})$ . Based on equ (2) and the components of the covariance matrix an index can be obtained. It can be considered as a measure of the information included by a lead in comparison with the leads that surrounds it, and is named *information index* (see equ. (4)). The information index is computed for every element of the covariance matrix and the position in the grid for which the value of the index is the greatest is considered as a PS position. The maximization of the information index is done iteratively. The first iteration selects as a first SP the lead for which  $C$  with a partition =1,  $n_{PE} = n-1$ , give a maximum value for the information index. The second iteration adds a second SP for which  $C$  has the partition  $n_{PS} = 2$ ,  $n_{PE} = n-2$ , and gives a maximum information index, a.s.o. The  $n_{PS}$ -th iteration, the last one, adds the  $n$ -th SP for which  $C$  with partition  $n_{PS, n_{PS}}$ ,  $n_{PE, PS} = n - n_{PS}$  maximizes the information index for the final form. In detail, the process will be described in the followings.

The information index for a position  $j$  is [1]:

$$I_j = \text{tr}(C_{11}^{-1} L_{11}) = \sum \sigma_i^2 \rho_{ij}^2 \quad (4)$$

where

$$C_{ij} = \rho_{ij} \sigma_i \sigma_j \quad (5)$$

$\rho_{ij}$  - correlation coefficient between position  $i$  and  $j$

$\sigma_i, \sigma_j$  - the standard deviation for the position  $i$  and  $j$  respectively.

Qualitatively, from the leads in a zone of the grid it will be selected the point from which the information "resembles" the least with the information from the leads in the neighbourhood. So, the position of the first SP is the one for which the dot product between the vector represented by the line (or column) with the same index in the covariance matrix with itself reported to its variance is the greatest. This lead is the least correlated with the one from its neighbourhood from the grid. In this case  $C_{11}$  and  $L_{11}$  are of  $1 \times 1$  dimension. In order to select the position for the second SP it is observed that the covariance matrix of the error, for  $m=1$ , is the covariance of the  $n-m$  ( $n-1$  for the second step) still not selected positions (reflected in  $C_{22}$ ), less a term that can be considered as the covariance corresponding to the first position selected previously.  $C_e$  (covariance of the error) will have the dimension  $(n-1) \times (n-1)$ . Considering the fact that the covariance matrix for the  $n-1$  not selected positions is  $C_e$  and the equ. (2)-(4) an algorithm can be developed in which is selected optimally (local), sequential, a SP position at one time. After a SP position was selected it stored and the information about him from the covariance matrix is eliminated. The error, after  $m$  steps is the same as in the case in which the selection of all  $m$  points was done in one step [Lux78]. This recommends the sequential algorithm as a simplest procedure.

The information from the heart level, reflected on the surface of the torso is found in all of the leads of the grid and so results a certain redundancy grade of the process and a certain correlation. This correlation implies the existence of an informational link between all the leads of the grid and can be quantified through a linear transformation,  $\hat{A}$  ( $n-m \times m$ ) (6) between the values measured in certain points, in particular the SP, and the values in the rest of the leads, the EP.

$$P_2 = \mathfrak{S} P_1 \quad (6)$$

$$\mathfrak{S} = \left\{ \mathfrak{S}^* \in \mathbb{R}^{n-m, m} \mid I_e = E[(P_2 - \hat{P}_2)^T (P_2 - \hat{P}_2)] = \min \right\} \quad (7)$$

$$\mathfrak{S} = C_{12}^T C_{11}^{-1} \quad (8)$$

With this, the values for the EP vector elements are:

$$\hat{P}_2 = C_{12}^T C_{11}^{-1} P_1 \quad (9)$$

The shortcoming of this algorithm is that it does not offer the guarantee of a global optimal set of SP. This shortcoming is overridden using the algorithm for GOSSP determination. A technique is **suggested** [1] through which from a lot of sets of SP is chosen **the optimal global set**. This technique is recommended for computers with high computing speed rates and can process great amounts of data. The number of SP will remain fixed (32).

The steps to follow in GOSSP algorithm are:

- application of the previously presented algorithm for the first grid point;
- computing of the index position  $p$  for the GOSSP and the transformation  $\mathcal{S}$ ;
- estimation, from the SP values the EP values;
- computation of the mean square error between the values of SP and EP;
- repeat of the previous 4 steps for all the remaining points;
- comparison of all the computed errors and selection of the least mean square error.

The **GOSSP** is the set of SP that has the least mean square error.

### 3.2 Program environment for GOSSP

Based on the theoretical developments presented we realised a program environment which yields the GOSSP. It was realised using the facilities given by the Visual C programming, mainly the MFC classes.

The loading of the file which contains the measured values for the potentials from the torso is done choosing from the menu **Fișiere (Files)** the option **Deschide (Open)** or with a mouse click on the corresponding command button from the toolbar. As a result, a window will open from which the file can be selected (with the extension .sgn) clicking **Deschide** (Fig. 2).



Figure 1. The main window of the application

In order to activate the SP option (Fig. 2) from **Fișiere** menu is selected the **Calcul T** option (Fig. 1) or the corresponding bitmap from the toolbar. The patient file and the number of the electrode are selected.

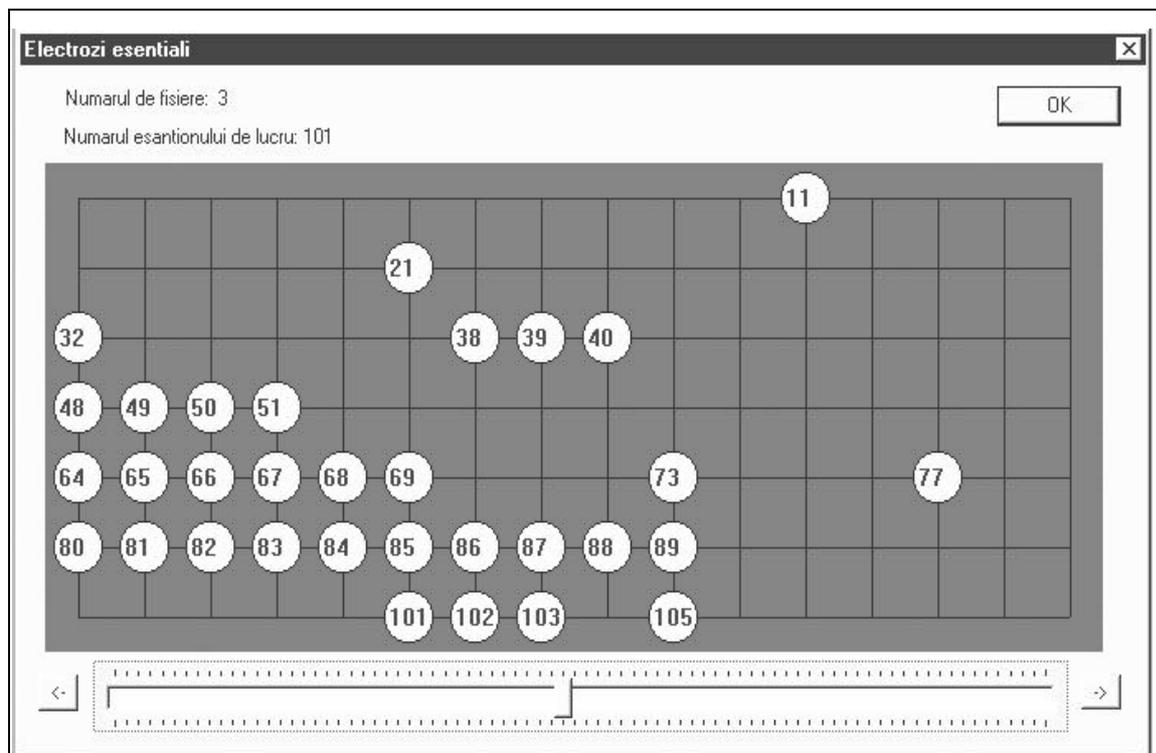


Figure 2. SP configuration

The EP values' computing is done from **Fiore** menu with **Estimare** option or by clicking the corresponding bitmap from the toolbar. The comparison is done visually: real/estimated maps.

#### 4 EXTENSION BY INTERPOLATION USING RIP. CASE STUDY

The measurements on the human torso are made using a waistcoat that surrounds it. On this waistcoat, as was suggested previously, are positioned the leads in which the signals from the torso are measured. They are placed in the nodes of a grid resulted from the intersection of longitudinal and transversal lines. The image from Fig. 2, which we name *cartesian grid*, is obtained if we unfold the waistcoat. On a perpendicular axis at the grid surface we will represent the potential magnitude. In this triorthogonal axis system, a point M will be characterised by three co-ordinates: M ( $x_M, y_M, z_M$ ).  $x_M$  and  $y_M$  are the coordinates from the cartesian grid plane and  $z_M$  is the associated potential from the torso. M describes a potential surface S. From the geometric point of view, the compression problem is stated as follows: determine a set of points  $\{M_{i,j}, i = 1, n_{PS}\}$  so that from the set  $\{M_i\}$  can be rebuilt the surface S.

This problem is similar with the one of RIP control [Dre96]. The set  $\{M_i\}$  of the SP has the role of the support points from the RIP method. The surface S is built going over two interpolation steps:

- The first step consists of *rough interpolation*, the interpolation method used is Shepard interpolation [She68]. It yields the set of the support points from  $n_{PS}$  to  $n$ , associating a support point  $M_j$  to every  $n - n_{PS} = n_{PE}$  point from the rest of the cartesian grid points.

The coordinates  $x_{M_j}, y_{M_j}$  corresponding to these points are known and  $z_{M_j}$  is determined with the relation:

$$z_{M_j} = \left( \sum_{i=1}^{n_{PS}} w_{ij} \right)^{-1} \cdot \left( \sum_{i=1}^{n_{PS}} w_{ij} \cdot z_i \right), \quad w_{ij}^{-2} = \sum_{i=1}^{n_{PS}} \left[ (x_{M_j} - x_{M_i})^2 + (y_{M_j} - y_{M_i})^2 \right]^{\frac{1}{2}}, \quad j = \overline{1, n_{PE}} \quad (10)$$

We consider  $\{M_k\} = \{\{M_i\}, \{M_j\}\}$  as the set of all the support points associated to the cartesian grid and  $\lambda \in [0, 10]$  an evaluation factor of the Euclidean distance  $(w_{ij})^{-1}$ .

The second interpolation step consists of a *fine interpolation* between the points of the set  $\{M_k\}$ . There can be used multilinear methods or other methods, depending of the density of the support points.

In section 2.2 there were presented the two steps followed in order to solve an application. The first step has, in the present case, a less elaborated character. In order to determine the SP is adopted the number  $n_{PS}$  and then is randomly tested the first degree of freedom and is compared, on statistical basis, in different configurations obtained. Three aspects must be retained:

- a.) For an adopted configuration of the SP the result (the values  $z_{M_j}$  from the nonsignificant points) depends on the value of the evaluation factor  $\lambda$ . Therefore the tests must take into account the second degree of freedom.
- b.) Covering the human torso, with an imaginary longitudinal line we will note the periodicity of the potentials. The display of the waistcoat like in Fig. 2 doesn't emphasise this periodicity. The shortcoming was eliminated considering a fictitious waistcoat that surrounds three times the torso. In this manner the configuration of the significant and nonsignificant points is tripled. As relation (10) states, the influence of the far points is reduced and we can consider that through this extension, for the waistcoat in the middle there are respected the periodicity conditions.
- c.) The application refers to the compression and reconstruction of a string of images. Following the complete results based on simultaneous processing of ECG signals from all the nodes of the cartesian grid during several cycles and representing the potential surfaces at different moments resulted from sampling with a constant step (hundreds of samples/period) we observed that the string of images is generally periodic and resembles with something "we can see from a satellite that surrounds a planet". Analysing such images we deduced that, generally, there are suited configurations of significant points uniformly distributed on the nodes of the cartesian grid.

The second step is reduced at the utilisation, whenever the case, of the information from the SP following successively the two interpolation steps.

For RIP method application we used a RIP development environment written in MATLAB based on a model from [3] from another research [5]. In Fig. 3 are suggested several elements and facilities of this environment. They exceed the necessities of the problems referred in the paper.

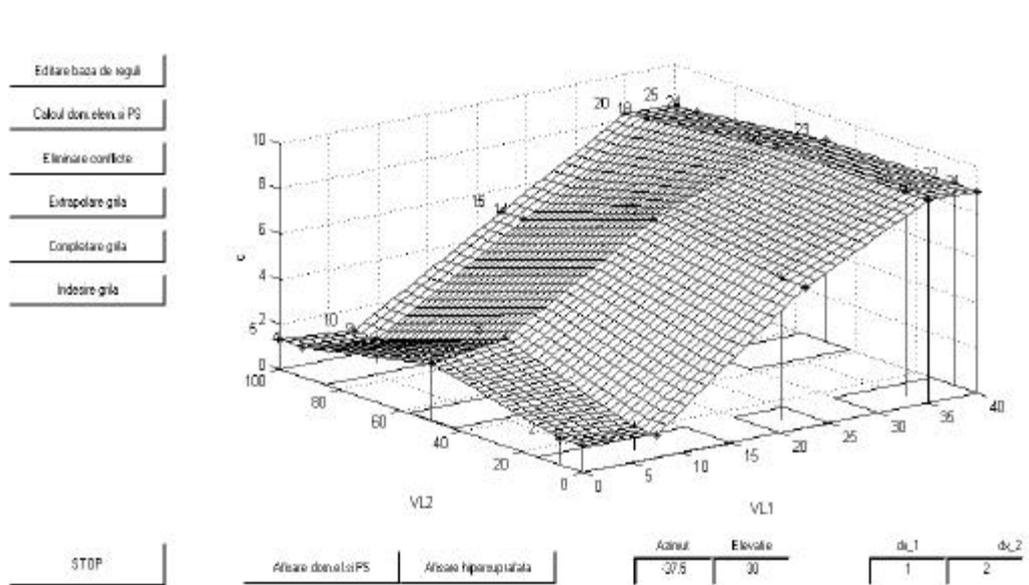


Figure 3. An image of the RIP development environment

## 5 COMPARISON OF RESULTS. CONCLUSIONS

The previously mentioned (in section 3 respectively 4) development environments for ECG data reduction and reconstruction, Ecg Manager and RIP, have the following features: provides the user with all the functions useful for SP selection and BSPM/PST representation and a user-friendly interface.

The environments were intensively used for the study of the usefulness of the compression methods presented above. The obtained results are promising. Qualitatively, physicians did the validation. The representations resulted as BSPM and PTS. The number of patients from the development set was reduced: 10. This matter affected especially the first method (BSPM) for which the SP configuration is better as the data amount is as greater as possible.

Fig. 4, 5, and 6 illustrate the types of images yielded by the environments. So, Fig. 4 contains the image of two BSPM, one resulted from measurements in 112 leads (down) and the other resulted from 32 measured points and 80 estimated values (up), 'Estimat' (estimated) respectively 'Real' (real). On the image is specified the number of the sample (110) of the cardiac cycle for which is drawn the map and the minimum and maximum value of the potentials.

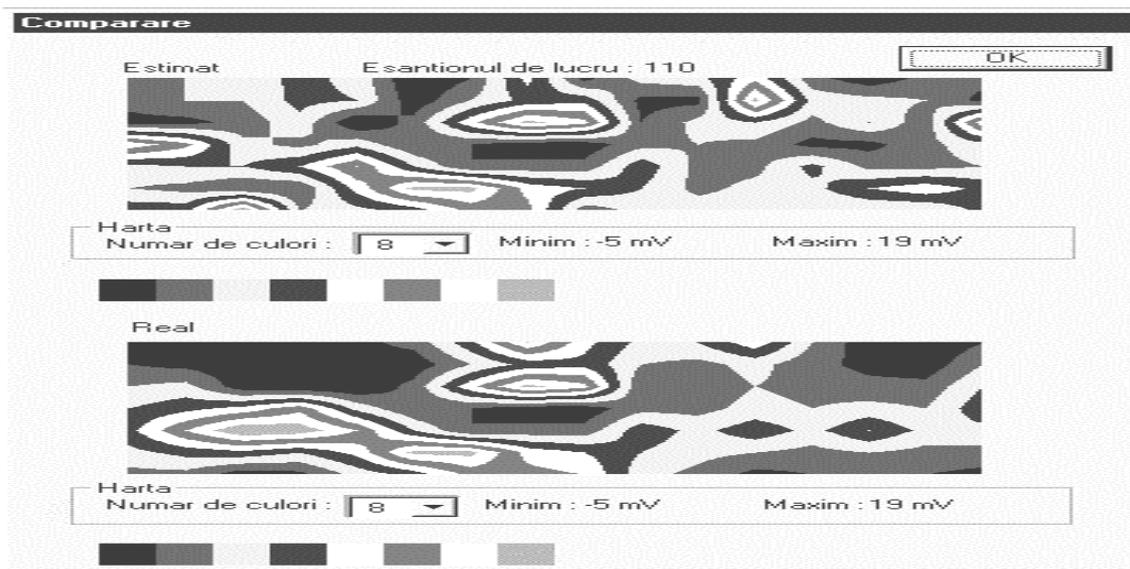


Figure 4. Estimated and real BSPM

Fig. 5 presents, comparatively, PTSs obtained after the application of the BSPM/PTS methods. In the upper image appears a "model-surface" (PTS model) that corresponds to the measurements in 112 leads. The other two are "image-surfaces" obtained through BSPM (in the middle) and RIP (down). The image-surfaces correspond to data reduction at half of the number of the grid leads (56), SP. The greater differences between the model surface and the first image-surface dew to a reduced number of patients.

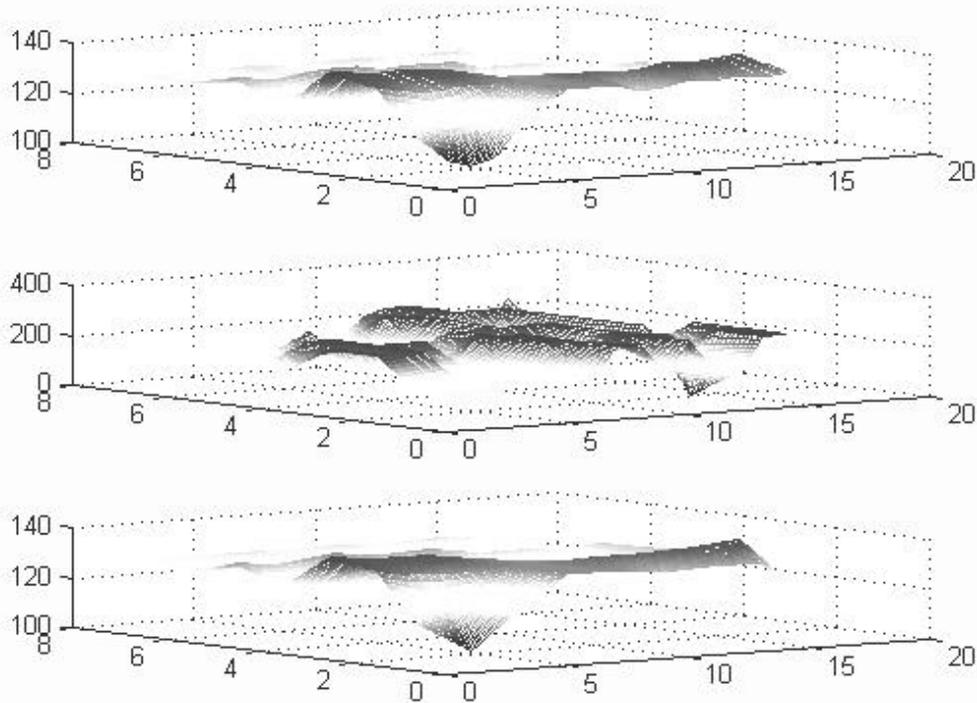


Figure 5. Model surface and image surface comparison for BSPM/PTS

Fig. 6 emphasises the fact that using RIP useful results are obtained compressing the number of leads at 42.

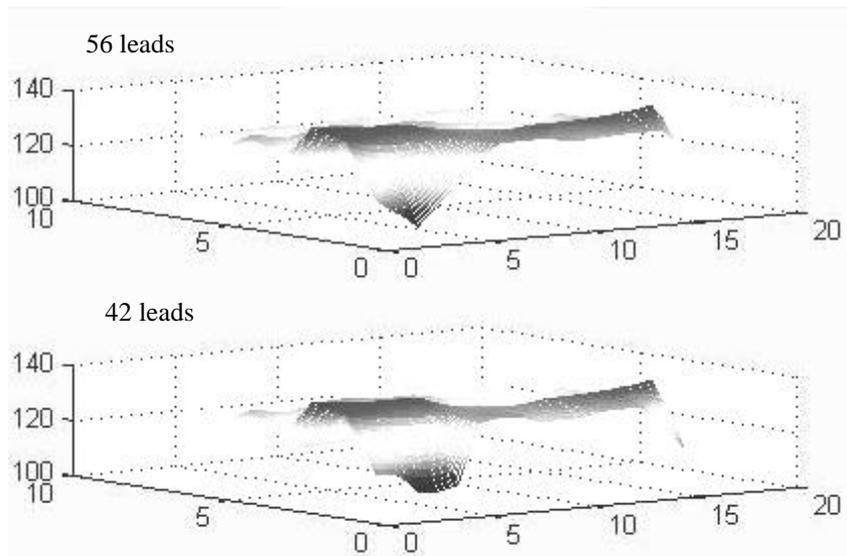


Figure 6. 56 vs. 42 comparison with RIP method

We appreciate that the obtained results are stimulated for continuing the studies. There are physicians who prefer diagnosis based on maps, respectively physicians who prefer diagnosis based on surfaces. The aspect of the surfaces from Fig. 5 and 6 depends also of the fine interpolation method used. In the present case we used bilinear interpolation. Using other fine interpolation

methods, smoother shapes can be obtained. The aspect is not very important because using consistently a certain type drives to habit with the types of images that it determines. Equally, there is no unanimity in what concerns the diagnosis based on static or dynamic (animation) images resulted from data compression.

## REFERENCES

- [1] Stoicu-Tivadar, L., "Aspecte sistemice i medii integrate asociate asistarii dignozei cardiace", Teza de doctorat, Facultatea de Automatica i Calculatoare, mai, 1999.
- [2] Lux, R. et al., "Limited Lead Selection for Estimation of Body Surface Potential Maps in Electrocardiography", IEEE Transactions on Biomedical Engineering, May, 1978.
- [3] Drechsel, D. : Regelbasierte Interpolation und Fuzzy Control, Vieweg-Verlag, 1996
- [4] Shepard, D. : A two-dimensional interpolation function for irregularly-spaced data, ACM National Conference, pg. 517-527, 1968
- [5] Dragomir, T.L., Mu, M. .a., Elaborarea unor strategii de control RIP i a unui mediu de dezvoltare adecvat. Studii preliminare referitoare la metoda RIP i la elaborarea unui mediu de dezvoltare, Grant CNCSIS Nr. 101, 1999.

AUTHOR(S): Department of Automation and Industrial Informatics, Faculty of Automation and Computers, "Politehnica" University of Timisoara, Bd. V. Parvan, Nr. 2, Timisoara, Romania, Phone: ++40 56 204333 ext. 200, E-mail: [dragomir@aut.utt.ro](mailto:dragomir@aut.utt.ro), [stoicu@aut.utt.ro](mailto:stoicu@aut.utt.ro).