Pattern Anal Applic
DOI 10.1007/s10044-006-0034-x

THEORETICAL ADVANCES

A semi-automatic system for segmentation of cardiac M-mode images

Luca Bertelli · Rita Cucchiara · Giovanni Paternostro · Andrea Prati

Received: 17 December 2004 / Accepted: 4 May 2006

Abstract Pixel classifiers are often adopted in pattern recognition as a suitable method for image segmentation. A common approach to the performance evaluation of classifier systems is based on the measurement of the classification errors and, at the same time, on the computational time. In general, multiclassifiers have proven to be more precise in the classification in many applications, but at the cost of a higher computational load. This paper analyzes different classifiers and proposes an evaluation of the classifiers in the case of semi-automatic processes with human interaction. Medical imaging is a typical application, where automatic or semi-automatic segmentation can be a valuable support to the diagnosis. The paper focuses on the segmentation of cardiac images of fruit flies (genetic model for analyzing human heart's diseases). Analysis is based on M-modes, that are gray-level images derived from mono-dimensional projections of the video frames on a line. Segmentation of the M-mode images is provided by classifiers and integrated in a multiclassifier. A neural network classifier, a Bayesian classifier, and a classifier based on hidden Markov chains are joined by means of a Behavior Knowledge Space fusion rule. The comparative evaluation is discussed in terms of both accuracy and required time, in which the time to correct the classifier errors by means of human intervention is also taken into account.

Keywords Performance evaluation · Neural network · Hidden Markov chains · Bayesian classifiers · Multiclassifier · Image segmentation · Cardiac imaging

L. Bertelli · G. Paternostro
The Burnham Institute, 10901 North Torrey Pines Road,
La Jolla, CA 92037, USA
e-mail: lbertelli@ece.ucsb.edu

G. Paternostro
e-mail: giovanni@burnham-inst.org

R. Cucchiara · L. Bertelli
Dipartimento di Ingegneria dell'Informazione,
University of Modena and Reggio Emilia,
Via Vignolese, 905, 41100 Modena, Italy
e-mail: cucchiara.rita@unimore.it

A. Prati (✉)
Dipartimento di Scienze e Metodi dell'Ingegneria,
University of Modena and Reggio Emilia,
Via Amendola, 2, 42100 Reggio Emilia, Italy
e-mail: prati.andrea@unimore.it

1 Introduction

Classifiers have been widely used in pattern recognition to segment images into homogenous regions. Possible applications range from SAR image analysis, to motion extraction, to medical imaging. Among possible classifiers, multiclassifiers have often demonstrated their superiority in terms of precision, since they exploit the advantages of multiple classifiers. However, it is also well known that fusing the results provided by multiple and different classifiers slows down the classification process w.r.t. single classifiers. However, this is not always true in the case of semi-automatic classification process: for certain applications, where the accuracy is of fundamental importance, the results provided by the automatic techniques must be further validated by an expert. In these cases, the evaluation of the time performance should consider also the expert's feedback and slower, but more precise, approaches can lead to the overall best

performance in terms of required time. With these premises, evaluating classifiers' performance must account for the best trade-off between precision and total computational time required to obtain ideal results with a semi-automatic approach.

In this paper, this analysis is carried out in the case of the segmentation of cardiac images. In particular, the analysis of heart images of the *Drosophila melanogaster* (common name: *fruit fly*) has been investigated, since it has become increasingly popular as animal *model* (i.e., simpler species with often similar genes) for the human heart [1, 2]. Heart disease is the main cause of death in developed countries and a large amount of the research in medicine is focused on the study of the human genome. Now that the human genome has been sequenced, the next challenge is to understand the function of the identified genes. Thus, *Drosophila* is one of the more studied models in genetics since its genome has been fully sequenced, it is a short-cycle species, its reproduction is very fast (a female starts laying eggs when it is 11 days old and it may lay up to 500 eggs in 10 days), and it is easy to culture and breed.

To investigate the effects on the fruit fly heart of life history, genetic mutations, and environmental conditions, some cardiac parameters must be monitored, and, in particular, the ratio between the minimum and maximum heart diameter in *end-systole* and *end-diastole*,[1] and the ratio between systolic and diastolic durations.

Since scientific considerations on models are significant only on a large population and manual measurements are neither precise nor fast, a (semi-) automatic methodology to collect and process data is desirable. Among possible sensors, a camera (in the visible spectrum) presents several advantages: it can collect much more information than other types of sensors; it does not require invasive interventions on the organisms, except for the immobilization of the organisms; it does not require long-lasting anesthesia of the organisms: only few seconds are enough for collecting the initial data to be processed to measure useful parameters. Moreover, it is reasonably cheap.

For these reasons, vision systems are often used to collect heart images to be analyzed. In particular, the first chamber of the heart of the fruit fly is filmed because it is the one with the largest wall movements and it is the closest to the glass slide and camera objective.

A widely-used approach for cardiac image analysis is to reduce complexity by using *M-modes*, which is the mono-dimensional projection of the image along a line (see Fig. 1).

This simplification permits a more accurate computation of the heart's parameters by reducing distractors, noise, and useless data. Once created, however, M-modes must be segmented to identify the border of the heart to detect end-systole and end-diastole. Unfortunately, as evidenced in Fig. 1, M-modes segmentation can be a challenge since illumination condition, fly abdomen's color, and acquisition noise can produce images with very blurred edges. In some cases, the distinction between points inside and outside the heart is very difficult (as in Fig. 1c).

The problem of heart segmentation can thus be viewed as a problem of two-classes classification of pixels. The classification of pixels in inner-outer part of the heart must be based on the gray level of the pixel and on the position of the pixel itself, possibly in a probabilistic framework. In our case, we compared three classifiers based on Bayesian framework, neural networks, and hidden Markov chains (HMC). In addition, in order to achieve high accuracy, a *multi-classifier* is proposed. A multiclassifier fuses the classification results from different classifiers by meaning of a suitable combining function, in order to take advantage of the best characteristics of each of them. We will describe the significant improvement in the results achieved by fusing the three classifiers.

The paper is structured as follows: the next section will report some related works in the field of cardiac imaging, image segmentation, and multiclassifiers; Sect. 3 will describe the M-mode creation and the preprocessing steps, while Sect. 4 will outline the classifiers used and the fusion function, discussing on the parameters used and their selection; then, Sect. 5 will show the experimental results and analyze them; finally, the conclusions will be drawn.

2 Related works

Cardiac images has been deeply analyzed in the past years, and different acquisition methods, segmentation techniques, and classification approaches have been proposed.

Most of the approaches to vision-based cardiac imaging reported in the literature are mainly based on 3-D imaging [3]. 3-D data are acquired with sophisticated technologies, such as cine magnetic resonance (MR) [4–7] and tagged MR [8–11], gated positron

[1] The term "end-systole" identifies the moment in which the heart has its minimum opening, whereas the term "end-diastole" identifies the moment with the maximum opening.

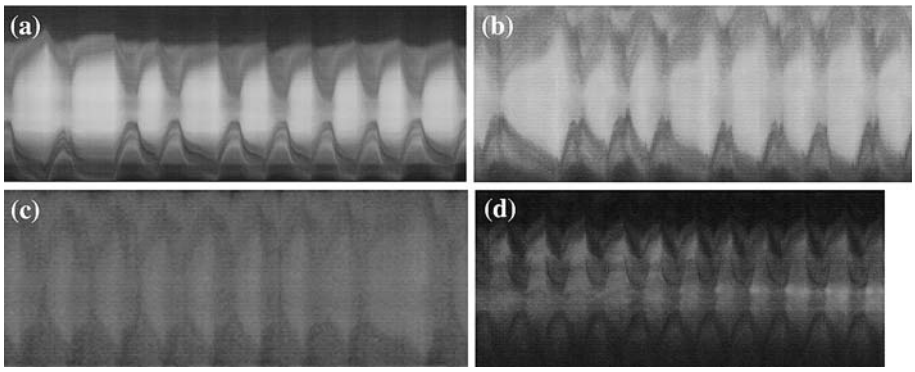


Fig. 1 Examples of M-modes

emission tomography (PET) [12] or single photon-emission tomography (SPECT) [13], echocardiography [14–17], or X-rays [18, 19]. The majority of cardiac imaging papers are related to the human heart, but there is also some noticeable work, such as that of Nahrendorf et al. [6] on rat and mouse heart. The advantages of the fruit fly as a model for heart disease have been described above (for example, the population in Nahrendorf et al.'s work consists of 8–26 rats, while in flies might be easier to obtain large sample numbers).

In our case, most of these acquiring techniques are not usable, for either obvious reasons of scale, or costs (too expensive), or possible alterations of the tests (affecting, for example, fruit fly heart function). Moreover, the information provided by these techniques are, often, not essential for our purposes: concentrating our efforts on computing the above-mentioned parameters, we might not need for information about 3-D of the heart or the tissue's properties. On the other hand, normal, visible-spectrum cameras are easy-to-install and easy-to-use, non-invasive, and they enable in vivo analysis of the heart.

Whichever acquisition system is used, an initial and fundamental step of cardiac imaging is the segmentation of the heart. The adopted techniques depend heavily on the dimensionality and the type of the input data, and on the final aim of the system.

In tagged MRI, the first step is usually the identification of the tags. In [9], a maximum-likelihood/maximum-a-posteriori (ML/MAP) method is proposed for myocardium identification. First, the tag is modeled with a Gaussian profile along the line perpendicular to the tag line. A set of tag line centers is estimated with a snake algorithm [20] based on a maximum-likelihood criterion. Ideally, this process should obtain for each of the tag centers within the myocardium to be on a tag line. Due to estimation errors, however, this is not always true. Thus, two MAP hypothesis tests are developed: the first determines if the center is inside the myocardium, the second if the center is part of a tag line. Finally, a pruning algorithm removes any center that does not meet a spatial-temporal continuity criterion. Osman et al. [11] exploits the harmonic phase of the tag patterns to extract a complex image (called *HARP*) that contains information not only for estimating the tag lines, but also for measuring motion and computing 2-D strain.

Tagging can simplify the identification of heart borders in MRI, but methods exist also in cine MR to obtain borders without tagging, such as that proposed by Staib and Duncan [21]. This method is based on probabilistic deformable global models of the contour shape. The parametric model is based on the elliptic Fourier decomposition of the boundary. Once borders have been identified, they can be exploited to build heart surface. For this purpose, tracking by using parametric B-spline surfaces has been proposed by Amini et al. [8] in tagged MRI, while McEachen and Duncan [4] proposed a quite general method based on the matching of local segments of the border in successive frames using a shape-based strategy. The final result of the matching procedure is a flow field that estimates heart movements. Optical flow was also exploited in gated PET by Klein and Huesman [12] to develop a 4-D deformable motion algorithm and estimate the motion.

Active contour or snakes [20, 22] are used very frequently to detect heart boundaries. This approach is based on the iterative fitting of a parametric curve on the image on the basis of internal and external energies that must be optimized. Where poor-quality data with low contrast are used, classical snakes tend to fail. Deformable templates [23] and trainable snakes [16, 24] have been proposed as solutions to this problem. In particular, Jacob et al. [16] applied snake to 2-D echocardiography that is probably the more challenging type of acquisition because it results in very noisy images and image information is highly anisotropic and position-dependent. Jacob et al. proposed a new method for tracking heart borders by means of dynamical tracking that combines a shape-space model

with a model describing the motion of the object in a Kalman-filter framework. These methods are not enough precise in the case of blurred edges.

Instead of following the heart boundaries in image sequences, cardiac imaging often exploits M-modes. M-modes [25] are often used in echocardiography for the analysis of left ventricular (LV) wall motion dynamics with the purpose of diagnosis of myocardial ischemia [26]. M-modes are created by projecting the heart's walls onto a line perpendicular to the walls. This process is often called in medical literature *anatomic M-mode* [27]. In such a manner the problem of border identification becomes an image segmentation problem, where classical approaches of image segmentation can be used. In the case of gray-level images (as the M-modes are) many methods have been proposed to segment them into meaningful areas (see for example the old but excellent surveys of Haralick and Shapiro [28], and Pal and Pal [29]): gray-level histograms, fuzzy sets [30], MRF models, neural networks, binarization and morphology [31], and many others. In [30], a method for histogram thresholding based on similarity between gray levels is proposed. The similarity is assessed through a fuzzy measure and the method works for both bimodal and multimodal histograms. In [31], instead, an anisotropic diffusion model for adaptive thresholding of bimodal images is presented. Differently from other proposals, the proposed diffusion model tries to smear edges rather than to preserve them.

Among these methods, since borders are not well defined and the gray level is not uniform, we propose to look at the segmentation task as a clustering/classification task: each pixel of the M-mode must be assigned to one of the two classes considered (inside and outside the heart) and traditional classifiers can be developed.

As stated above, we designed a *multiclassifier* that has the purpose of exploiting the strengths of different classifiers, bypassing their weaknesses [32]. Even if multilearner systems have a quite long and complex history, the application of multiclassifiers in the pattern recognition field has just begun in the last 10 years and their qualities have not been completely explored yet. Kittler et al. in [33] have developed a common theoretical framework for combining classifiers which use distinct pattern representations and show that many existing schemes can be considered as special cases of compound classification where all the pattern representations are used jointly to make a decision. Similarly, in [34] the local accuracy of individual classifiers in small regions of feature space surrounding an unknown test sample is used as an estimate to combine the classifiers.

Without going too much into details, multiclassifiers are used in pattern recognition with success in the recognition of handwriting cursive words [35], in decision making in ballistic missile defense applications [36], in the analysis of remote sensing data [37], and in fingerprint classification [38]. Fusion of different classifier is, however, useful only if they are mutually complementary, i.e., they make different kinds of mistakes. In the literature there are different techniques to design a set of complementary classifier [32], suitable for a mixture of classifiers of the same kind (for example a set of N neural networks). However, in our work, we built up separately three different classifiers with structures, parameters and features different among each other, and tests demonstrate their complementary in pixel assignment.

Another key aspect in a multiclassifier system is the choice of the fusion method. Methods for fusing multiple classifiers can be distinguished according to the type of information produced by the individual classifiers [32]: *abstract-level* outputs, *rank-level* outputs, or *measurement-level* outputs. For each of the above categories, methods can be further subdivided into ''with fixed rules'' and ''with trained rules''. In the abstract-level type, each classifier outputs a unique class label for each input pattern, which is called ''the opinion''. An example of simple fixed rule in the case of abstract-level outputs is that of deciding in accordance with the majority of the opinions. An example of trained rules is the Behavior Knowledge Space (BKS) [32] which is the method adopted in our multiclassifier and it will be explained in detail.

3 Pre-processing steps

Analyzing in details the heart movements is a problem that should be tackled in four dimensions (three spatial dimensions plus time). For the detection of the end-systole and end-diastole, however, the problem's complexity can be reduced. First of all, for these measurements, a high-speed camera has been used to record at a speed of 250 frames/s (Motionscope, Redlake MASD) at a resolution of 480×400, connected to a Nikon Diaphot-TMD inverted microscope. Using a standard camera the spatial dimensionality is automatically reduced from 3-D to 2-D. Then, the problem can be further simplified with the *M-modes* construction. A certain line (called *projection line*) in the first image of the sequence is defined. This line should be perpendicular to the heart walls (i.e., it should coincide to the minor axis of the heart), it should be long enough to cover the heart walls in their maximum

opening (it is good practice that the line is slightly longer than the case of end-diastole), it should be approximately centered with respect to the major axis of the heart, and it is kept constant all over the sequence.

Since the selection of the projection line affects heavily the overall performance, manual selection by an expert user is preferable.

As shown in Fig. 2, the M-mode is created by reporting the gray level of the pixels of the input frame n belonging to the projection line as the n-th column of the M-mode image. Thus, the resulting image has a size of $M \times T$, where M is the length (in pixel) of the projection line and T is the number of frames of the sequence and represents the time. It is clear that the M-mode image presents two different regions: the first one is composed by the points belonging to the innermost part of the heart and it is located close to the center of the M-mode; the second one is made by the points belonging to the outside part of the heart and is located near the horizontal bounds of the image. The former region is brighter because during the contraction the fruit fly heart is periodically filled with a blood-like substance (called *hemolymph*) that has a high reflection index: during the end-diastole the heart is wide opened and, thus, it contains more hemolymph.

Before starting to segment the M-mode image, a pre-processing step is convenient to reduce useless artifacts. In fact, the fruit fly heart is a symmetric structure, but, due to technical problems during the acquisition phase (for example, it is almost impossible to put the back of the fly exactly parallel to the glass slide under the microscope), during the movie stream the heart does not present anymore symmetric features. Therefore, the created M-mode image does not represent in the best way possible the dynamic behavior of the heart (an example is reported in Fig. 3a).

Since the goal is to extract precise end-systoles and end-diastoles, and other information can be neglected, a normalization procedure in which the M-mode is flipped with respect to its ideal middle line is performed. Being y_0 the y coordinate of the middle line and $I(x, y)$ the gray level value, for each point, $I(x, y)$ is substituted by the value $(1/2)(I(x, y) + I(x, 2y_0 - y))$, i.e., the mean between its gray value and that of the point at the same distance from the middle line. The result for the M-mode of Fig. 3a is shown in Fig. 3b. Obviously, with this process the sharpness of the border can be reduced, but the fuzzy walls are better distinguishable.

Once the image is enhanced, main process consists in the point classification into innermost and outermost parts of the heart. The details on the segmentation based on the classifiers and multiclassifier will be reported in the next section. As it will be explained, however, all the classifiers used need some a-priori assumptions. An initial shape that approximates very roughly the border

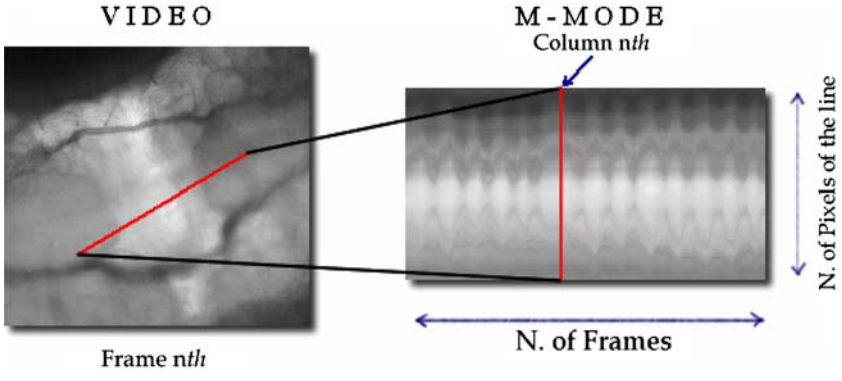


Fig. 2 M-mode creation process

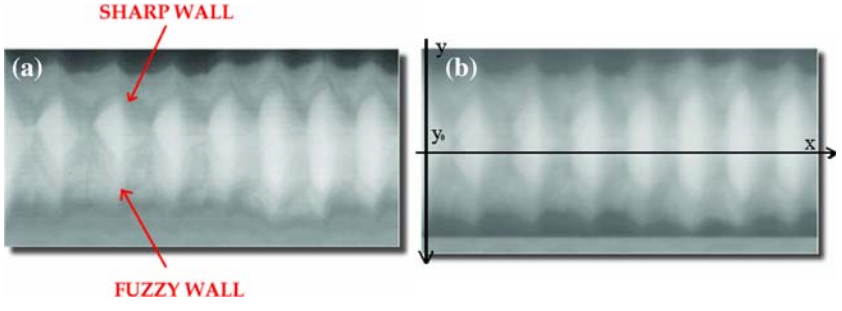


Fig. 3 Example of a asymmetrical M-mode **a** before the pre-processing and **b** after the flipping

delimiting the two regions of the M-modes will be used as training set to extract a-priori statistics. The initialization is obtained by asking to the user to select into the M-mode four points corresponding, approximately, to the top and bottom limits for the first end-diastole and end-systole (see Fig. 4). The initialization is obtained by the four points provided by the user copying the initial shape along the x-axis.

4 Segmentation with classifiers

The next three subsections will provide details on the three classifiers used in this analysis, while the last section will discuss on the fusion function. We will mix three classifiers: a neural network classifier (NNC), a Bayesian classifier (BAC), and a classifier based on HMC. From the tests they are complementary in segmenting M-modes, and thus they are ideal for a multiclassifier system.

As a basic notation for the following sections, we will refer to ω_{IN} for the class containing the pixels of the innermost part of the heart and to ω_{OUT} for the class of the outermost part. Moreover, F will indicate the feature vector used for the classification.

4.1 Neural network classifier

Neural networks have been widely studied in the last decades and their application ranges from pattern classification [e.g., for optical character recognition (OCR) or fingerprint classification] as in [39] or [40], to noise reduction (e.g., to recognize patterns corrupted by noise in voice or images) as in [41], to prediction (e.g., to predict the value of a variable given historical values—loads, market, weather, etc.) as in [42], to image segmentation as in [43] or [44], to many others.

Neural networks are computational systems composed of many simple processing units, which communicate with the help of a large number of weighted connections. The original model took inspiration from the nature of the human brain. Like human brains, artificial neural networks consist of processing units, artificial neuron, also called *nodes*, and weighted connections, also known as *links*. Neural Networks are parallel distributed models that have several distinguishing features:

1. A set of processing units;
2. A combination function that combines the input values;
3. An activation function, which determines the new level of activation based on the actual inputs and the bias;
4. An activation state for each unit, which is equivalent to the output of the unit;
5. An external input (bias, offset) for each unit;
6. Connections between the units. Generally each connection is defined by a weight w_{jk} that determines the effect that the signal unit j has on unit k;
7. A method for information gathering (learning rule).

Each neuron j has his own input signals (one or more) x_{j0}, x_{j1}, ..., x_{jn}, but only one output y_j The *combination function* is set as $a_j = \sum_{i=1}^{n} w_{ji}x_i + \theta_j$, where θ_j is called *bias* term. The training algorithm adopted is the *Back-Propagation* algorithm and the activation $y_j = \phi(a_j)$ is provided by the *sigmoid* function, i.e., $\phi(a) = \frac{1}{1+e^{-ca}}$, with c a suitably defined constant.

There are many different types of NN used in the classification literature. Since our inputs are not binary and the separation function is not linear, a multi-layer perceptron (MLP) should be the best suitable choice [45]. MLP is a layered feed-forward network with one or more hidden layers. An extensive study of the performance has been carried out to select the best structure of the network in terms of number of hidden layers, number of hidden units, activation function, parameters of the training algorithm, etc.

The first problem is to select the features to code as inputs of the network. For our purposes, the gray level of the point is not enough and the distance from the center of the M-mode is added. Even if not required by the MLP, it is often convenient to normalize the inputs to the network in order to feed the MLP with real number whose values range from 0 to 1. This will

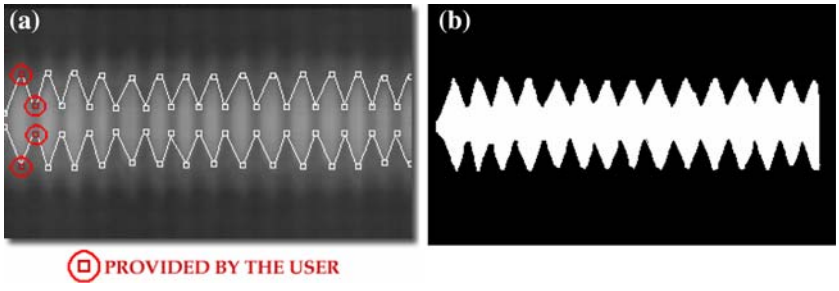


Fig. 4 Initialization process (**a**) and resulting mask image (**b**)

equalize the contribution of each input to the network. To achieve this, given $I(x, y)$ the gray level of the point p at coordinates (x, y) and $D(p) = |y - y_0|$ its distance from the center y_0 of the M-mode, the inputs to our MLP are:

$$F = [f_1, f_2]; f_1 = \frac{I(x,y) - \min \mathrm{GL}}{\max \mathrm{GL} - \min \mathrm{GL}}; f_2 = \min\left(\frac{D(p)}{y_0}, 1\right), \quad (1)$$

where min GL and max GL are the minimum and the maximum gray level, respectively.

Essential parameters of a MLP are the number of hidden layers and the number of hidden units. We compared the classification results for 11 different configurations, considering that it has been demonstrated that a two hidden layer net can represent an arbitrary decision bound (thus, there is no need to investigate on more than two hidden layers). The results are reported in Table 1, that reports the average error with respect to the ground-truthed results (obtained by manually segmenting the M-modes) and the time required to train the network with 100 training generations (on a standard PC with a Pentium 4 processor at 2.4 GHz and 512 MB of RAM).

From Table 1 it is easy to see that the networks with two hidden layers outperform those with only one layer. Moreover, increasing the size of the network, in this case the number of hidden units, does not always mean improving the performance. The size increase affects two aspects:

- It increases the time required by the training and so the whole system gradually slows down;
- The *overfitting* problem arises: the network has too much information processing capacity compared to the limited amount of information contained in the training set, and, thus, cannot be correctly trained.

Table 1 Simulation results for one and two hidden layers and different numbers of hidden units

Network topology	Average error	Average training time (s)
1 HL 3 nodes	0.307987	16.4
1 HL 6 nodes	0.200513	24.8
1 HL 9 nodes	0.364186	32.2
1 HL 15 nodes	0.362741	50.4
2 HL 2–3 nodes	0.075220	20.2
2 HL 4–5 nodes	0.071217	34.7
2 HL 6–8 nodes	0.080274	57.2
2 HL 8–10 nodes	0.068429	76.5
2 HL 10–14 nodes	0.074513	121.2
2 HL 16–20 nodes	0.201803	189.7
2 HL 30–40 nodes	0.747923	539.0

Looking at the average error in Table 1 and after a visual feedback provided by medical staff, we developed a network with two hidden layers with 8 and 10 units, respectively, and with 2 input units. Training data are provided by the initialization described in Sect. 3.

Another parameter that must be considered is the *learning rate* η. Results achieved with different values of η are reported in Table 2. This parameter enables to control the speed of the learning: the smaller it is, the smaller the changes to the weights will be, the smoother will be the trajectory in the weight space but the slower the network will be; on the other hand, large η values can speed up the learning but the network may become instable.

We made the choice to select the learning rate that corresponds to the lower average error ($\eta = 0.05$), even if this can slow down the training process. However, to reduce the computational time, we perform the training only once and off-line using a wide collection of M-modes as training set. The achieved results have proven to be comparable with those reachable with on-line training.

4.2 Bayesian classifier (BAC)

This classifier is very simple but frequently used due to its easiness of implementation and to the good results that it can provide in many problems describable with a statistical model [46, 47]. The aim in the analyzed case is to associate all the pixels of the M-mode to either ω_{IN} or ω_{OUT} Thus, the Bayesian approach aims at finding the a-posteriori probability as:

$$P(\omega_{\mathrm{IN}}|F) = \frac{P(F|\omega_{\mathrm{IN}})P(\omega_{\mathrm{IN}})}{P(F)}, \quad (2)$$

where F is the feature vector. Similarly, we can define $P(\omega_{\mathrm{OUT}}|F)$. The problem is formulated as in a MAP framework and the pixel will be associated to the class with the highest probability. The a-priori terms $P(\omega_{\mathrm{IN}})$ and $P(\omega_{\mathrm{OUT}})$ are computed from the initial mask as in Fig. 4 averaged on a set of training M-modes.

Table 2 Simulation results for different values of η

η coefficient	Training steps	Average error
0.025	100	7.294690
0.050	100	7.142626
0.100	100	7.824680
0.500	100	19.629180

To improve the classification, too rough if based only on the gray level of the points, geometrical information should be considered. Given the closed curve delimiting the mask image constructed in the initialization, the more distant a pixel is from the border, the higher is the probability to be either inside or outside the heart. On the other hand, for the pixels close to the border the only available information is the intensity.

To take this into account, a function $f(y_i)$ (Fig. 5) is used. This function value ranges from 0 to 1 and it represents the dominance of the intensity information on the geometrical information. Moreover, it is created to grow in a non-linear way.

Referring to Fig. 5, the function $f(y_i)$ can be written as:

$$f(y_i) = \begin{cases} \frac{y_i^2}{y_1^2} & 0 \le y_i \le y_1 \\ 1 - \left| \frac{y_i^2 - 2y_i y_0 + y_1^2 + 2y_1(y_2 - y_1)}{(y_2 - y_1)^2} \right| & y_1 \le y_i \le y_2 \\ 1 - \left| \frac{-y_i^2 + 2y_i y_M - y_2 y_M}{(y_M - y_2)^2} \right| & y_2 \le y_i \le y_M \end{cases} \quad (3)$$

Please note that y_1 and y_2 are actually function of x: for a given x_i, the values $y_1 = y_1(x_i)$ and $y_2 = y_2(x_i)$ are obtained by the mask image provided by the initialization.

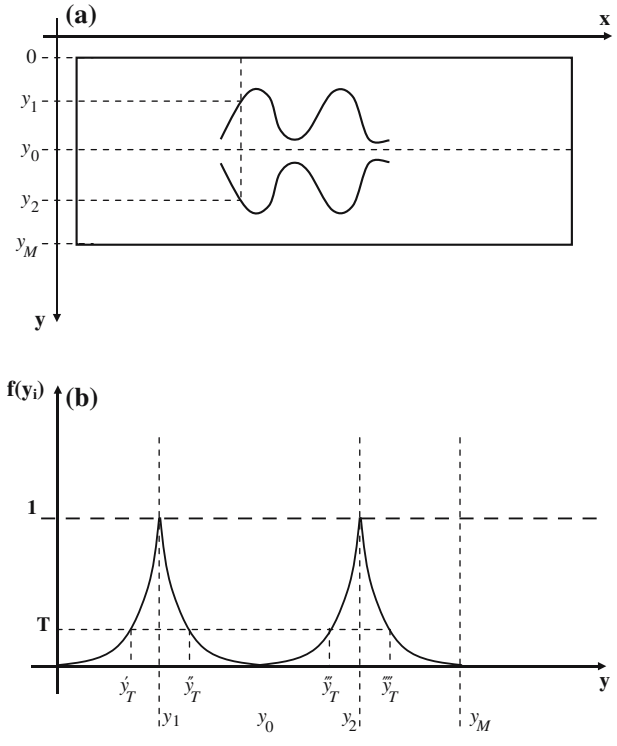


Fig. 5 Geometrical parameters calculation

Thus, for a point (x_i, y_i) having the gray level I_i, we assume that $P(F|\omega_{\mathrm{IN}}) = P(I(x,y) = I_i, x = x_i, y = y_i|\omega_{\mathrm{IN}}) = P(I_i, x_i, y_i|\omega_{\mathrm{IN}})$. With the strong assumption to consider gray-level and spatial information as independent, this equation can be written as:

$$P(I_i, x_i, y_i|\omega_{\mathrm{IN}}) = P(I_i|\omega_{\mathrm{IN}}) \times P(x_i, y_i|\omega_{\mathrm{IN}}). \quad (4)$$

$P(x_i, y_i|\omega_{\mathrm{IN}})$ is the probability for a point of having the coordinate (x_i, y_i), given to belong to the class ω_{IN} Using the function f reported in Fig. 5b, the probability $P(I_i, x_i, y_i|\omega_{\mathrm{IN}})$ can be written as:

$$P(I_i, x_i, y_i|\omega_{\mathrm{IN}}) = \begin{cases} 0 & 0 \le y_i \le y'_T \\ P(I_i|\omega_{\mathrm{IN}}) * f(y_i) & y'_T \le y_i \le y''_T \\ 1 & y''_T \le y_i \le y'''_T \\ P(I_i|\omega_{\mathrm{IN}}) * f(y_i) & y'''_T \le y_i \le y''''_T \\ 0 & y''''_T \le y_i \le y_M \end{cases} \quad (5)$$

where y'_T, y''_T, y'''_T and y''''_T are the y-coordinate in which the function takes the value of a threshold T. This threshold identifies where the geometrical information are sufficient reliable to be sufficient for the classification: the higher the threshold, the more reliable the geometrical information are considered, the tighter the cusp is around the borders (y_1 and y_2 coordinates).

In the case of $P(I_i, x_i, y_i|\omega_{\mathrm{OUT}})$, Eq. (5) becomes:

$$P(I_i, x_i, y_i|\omega_{\mathrm{OUT}}) = \begin{cases} 1 & 0 \le y_i \le y'_T \\ P(I_i|\omega_{\mathrm{OUT}}) * f(y_i) & y'_T \le y_i \le y''_T \\ 0 & y''_T \le y_i \le y'''_T \\ P(I_i|\omega_{\mathrm{OUT}}) * f(y_i) & y'''_T \le y_i \le y''''_T \\ 1 & y''''_T \le y_i \le y_M \end{cases} \quad (6)$$

We will demonstrate that the final segmentation is very good, comparable to other more complex classifiers.

4.3 Hidden Markov chain classifier

The field of application of hidden Markov models is extremely vast. Among this family of models, monodimensional HMM (also known as HMC) are the most frequently used. We can easily find several examples of application in the pattern recognition and image processing area: image segmentation [48–50], handwritten word recognition [51], speech recognition [52], noise restoration [53].

HMC have been chosen because, differently from neural networks or Bayesian approach, consider also

the features of the pixel's neighbors and the previous decisions in the chain. HMC can be adapted to 2-D analysis (necessary for images) through a linear scan of the image. This simple solution has found to be as effective as other more sophisticated approaches, such as the Hilbert-Peano [54] scan.

A HMC can be mathematically defined as HMC = {*S, V, A, B,* Π} where $S = \{s_1, ..., s_M\}$ is the set of M possible states, $V = \{v_1, ..., v_L\}$ is the output alphabet, $A = \{a_{ij}\}_{ij-1}...M$ is the set of transition probabilities, $B = \{b_{jk}\}$, $j = 1...M$, $k = 1...L$ is the set of output probabilities, and $\Pi = \{\pi_1...\pi_M\}$ is the set of initial state probabilities. Over time, the HMC will output an observation sequence O while visiting the states in the sequence Q. It also holds that:

$$\begin{aligned} a_{ij} &= P\big(s_i(t+1)|s_j(t)\big), \\ b_{jk} &= P\big(v_k(t)|s_j(t)\big). \end{aligned} \tag{7}$$

In our particular case, the HMC consist of two states (one corresponding to the ω_{IN} class, the other to the ω_{OUT} class) and 256 observations (corresponding to the 256 gray levels that the pixel can assume in the M-mode image). First, the HMC has to be trained to determine model parameters, i.e., the state transition probabilities and the output probabilities. The state transition probabilities are determined by counting the number of transitions of a certain type (from state 0—inside to state 1—outside, for example) that occur in the initial shape described in Sect. 3. The output probabilities, instead, are computed by evaluating on the initial shape the value of the observation (the gray level) in dependence of the state. Figure 6 shows a graph with an example of observation probabilities.

The aim of the classifier is to find the most probable sequence of hidden states, given a sequence of M-modes points (i.e., a sequence of observations), and given the state transition and output probabilities. This is also known as the *decoding problem* of the evaluation of HMC and one of the most popular techniques to solve it is the Viterbi algorithm [55–57].

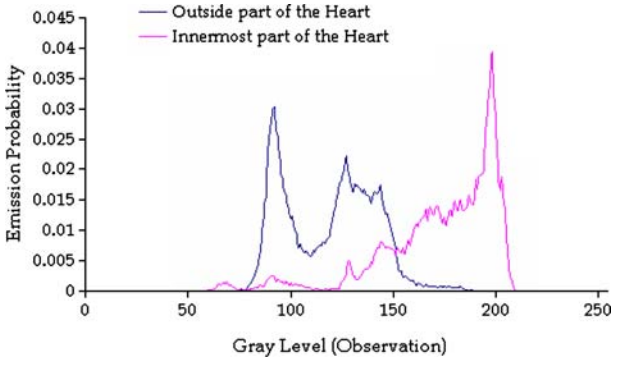


Fig. 6 An example of observation probability distribution

The characteristic of this classifier to consider also neighbors' features would help to decide correctly especially in the part of the M-mode with the edge between the two regions. Figure 7 reports an example of the result of the classification, expressed by the border traced by the HMC classifier.

4.4 The fusion function: the BKS algorithm

As already stated in Sect. 2, to fuse the outputs from the three classifiers this system uses an abstract level method implemented with trained rules, called BKS [58, 59]. The BKS tries to estimate the a-posteriori probabilities from a training set, by computing the frequency of each class corresponding to each combination of the classifiers decisions. Let $S_1, S_2, ..., S_N$, be the decisions of the N classifiers and c_i each of the K classes to be evaluated. The a-posteriori probability for class c_i is defined as follows:

$$P(c_i|S_1, S_2, \ldots, S_N) = \frac{N(c_i|S_1, S_2, \ldots, S_N)}{\sum_{j=1}^{K} N\big(c_j|S_1, S_2, \ldots, S_N\big)}, \tag{8}$$

where $N(.)$ indicates the occurrence of the correct classification of c_i given the combination of the outputs $S_1, S_2, ..., S_N$. The occurrences are computed on the training set by using the corresponding initial shapes.

5 Experimental results and discussion

By means of the algorithms described in the previous Sections, the border between the innermost part of the heart (class ω_{IN}) and the outermost one (class ω_{OUT}) can be extracted. Let $N_{\omega_{\mathrm{IN}}}(i)$ be the number of points assigned to the class ω_{IN} in the i-th column of the M-mode. By finding the local minima and maxima of the function $N_{\omega_{\mathrm{IN}}}(i)$, we can identify the end-systoles and the end-diastoles, respectively. Let n be the number of beats (corresponding to as many end-systoles and end-diastoles), we call N_{S_j} and N_{D_j} the corresponding value of $N_{\omega_{\mathrm{IN}}}(i)$ for the j-th beat. Thus, the

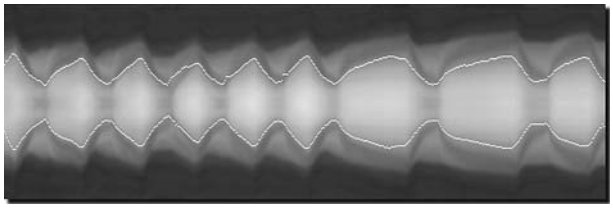


Fig. 7 An example of borders found with the HMC classifier

average diameter amplitude in end-systole can be computed as follows:

$$N_{S_{\text{mean}}} = \frac{\sum_{j=1}^{n} N_{S_j}}{n}. \tag{9}$$

Similarly, we obtain the value $N_{D_{\text{mean}}}$ in the case of end-diastole.

With the same approach, it is computed the average ratio between the end-systole and end-diastole diameters ξ_A:

$$\xi_A = \left(\frac{N_S}{N_D}\right)_{\text{mean}} = \sum_{j=1}^{n} \left(\frac{N_{S_j}}{N_{D_j}}\right) \Big/ n. \tag{10}$$

The other important parameter, the ratio between the systolic and diastolic durations ξ_T, is computed as follows:

$$\xi_T = \left(\frac{TS}{TD}\right)_{\text{mean}} = \frac{\sum_{j=2}^{n} \left(\frac{Ts_j - Ts_{j-1}}{Td_j - Td_{j-1}}\right)}{n}, \tag{11}$$

where Ts_j and Td_j are, respectively, the instant in which the j-th end-systole and the j-th end-diastole occur.

In the reminder of this section we will focus our discussion only on $N_{D_{\text{mean}}}$, $N_{S_{\text{mean}}}$, and ξ_A, and not on the parameter ξ_T described in Eq. (11). The reason why we do not consider the parameters on time, such as the durations of diastolic/systolic phase, was the fact that the temporal parameters are less sensitive to a good segmentation, thus are less significant for our purposes.

5.1 Performance of the automatic process

Usually, ground-truthed data are used to compare the performance of different approaches. In our application, since it is quite hard to say exactly where the heart walls are, the ground truth created by manually scratch could be affected by errors. An alternative approach is to perform an a-posteriori validation of the results provided by the automatic technique. In practice, an expert evaluates the classifier output, correcting by hand the misclassified points, and, at the end, calculates the difference between the actual output and the desired one. As a consequence, in order to evaluate the goodness of the classifier, the precision is indirectly taken into account, since it affects the overall required time. In particular, two important parameters can be evaluated: the *number of points added* by the user (when the system missed either an end-diastole or an end-systole) and the *number of points changed* by the user (when the system caught correctly the end-diastole/end-systole, but placed the diameters border in an incorrect position). The more precise the automatic classification is, the less the time to spend in correcting its results. The other parameter corresponding to the over-segmentation of the M-modes (i.e., the number of points *deleted* by the user) has not been considered since the proposed classifiers tend to create more false negatives than false positives.

For each of these parameters and for each classifier implemented, we repeated the measurements over a set of 15 M-modes, created with about 500 frames each. Moreover, to reduce the measurement noise and to obtain more reliable data, we repeated five times each measurement.

The summary of the results are reported in Table 3. Each of the first three main rows reports the average error w.r.t. the version manually corrected. Numbers in brackets represent the corresponding standard deviations. Fourth and fifth rows show the average number of points added and changed by the user, respectively. Finally, last row reports the required computational time (in seconds) on the same computer of Sect. 4.1. Moreover, Fig. 8 shows the visual comparison of the segmentation of single classifiers w.r.t. the multiclassifier. NNC and BAC results are inconvertibly incorrect (NNC tends to heavily under-segment the inner part, while BAC make some slight but evident errors). HMC

Table 3 Comparison of the performance of the neural network classifier (NNC), the Bayesian classifier (BAC), the hidden Markov chain classifier (HMC), and the proposed multiclassifier (MCS)

Parameters	NNC	BAC	HMC	MCS
$N_{D_{\text{mean}}}$	3.45(0.81)	1.82(0.49)	1.64(0.36)	0.64(0.16)
$N_{S_{\text{mean}}}$	2.07(0.92)	2.38(0.81)	2.19(1.31)	0.78(0.44)
ξ	0.04(0.012)	0.041(0.013)	0.036(0.018)	0.015(0.0070)
Points added	0(0)	0.19(0.11)	0.067(0.067)	0(0)
Points changed	3.57(0.74)	3.45(0.53)	2.79(0.55)	1.19(0.35)
Time (s) T_{A}	13	6.333	8.533	16.067

Figures in the first three main rows are the average error. In brackets the standard deviations

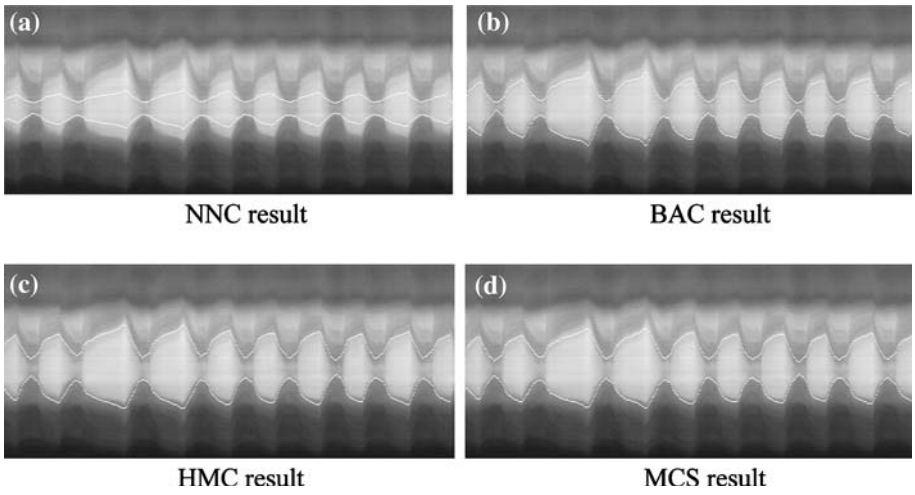


Fig. 8 Visual comparison of the segmentation results

and MCS results are comparable, but the latter slightly outperforms the former (look, for instance, to the end-diastole on the third beat).

5.2 Performance of the semi-automatic process

From Table 3 and Fig. 8, it is evident that MCS outperforms the single classifiers in terms of classification accuracy. It is also clear that the MCS requires more computational time to obtain results, though the time required is smaller than the mathematical addition of the three average times required by the single classifiers (due to fixed times that can be shared among the classifiers). Using a multiclassifier also requires to train more classifiers, and this can be a long lasting process. It must be done, however, only once. Moreover, it is worth noting that the final objective of this system is to collect data with high accuracy. Ideally, manual computation of cardiac parameters should be adopted. Since this is unaffordable, the proposed approach is a good trade-off between manual (thus precise, but tedious) and automatic (thus, fast, but often imprecise) segmentation.

Actually, the required time can be compared in three cases:

- Complete manual detection of systolic and diastolic points in the M-mode (required time: T_M);
- Semi-automatic process in which the user intervenes to correct the points misclassified by the automatic system, as described above (required time: T_{SA});
- Complete automatic process, as in Table 3 (required time: T_{A}).

To compute the time of manual detection, an expert user has been asked to manually segment all the 15 M-modes of our test set. An average time T_{M} of 104 s and an average time for each couple of points T_{MC} of 4.949 s have been obtained.

The time required in the semi-automatic process is:

$$T_{\mathrm{SA}} = T_{\mathrm{A}} + T_{\mathrm{VALIDATION}} + T_{\mathrm{MC}} \times N_{\mathrm{CC}}\,, \tag{12}$$

where $T_{\mathrm{VALIDATION}}$ is a fixed time utilized by the user to analyze the results and to figure out which points have to be changed (we estimate this time about 3 s), and N_{CC} is the number of couples changed or added. We can compute the speed-up achieved with semi-automatic and automatic process as the ratios $\mathrm{SP_{SA}} = \frac{T_{\mathrm{M}}}{T_{\mathrm{SA}}}$ and $\mathrm{SP_A} = \frac{T_{\mathrm{M}}}{T_{\mathrm{A}}}$. Table 4 reports the comparison results.

This temporal analysis highlights the classifier which gives the best speed-up in the complete automatic case is the Bayesian classifier. It is the fastest but not as accurate as, for example, the multiclassifier system. In fact, if we consider the time required for the correction of the misclassifications, we can notice how the best speed-up is given by the multiclassifier where the human interaction is less frequent compared to the other classifiers.

6 Conclusions

This paper presents an approach to cardiac image segmentation based on a multiclassifier and an evalu-

Table 4 Comparison of the speed-up in automatic and semi-automatic processes

Classifier	T_{A} (s)	$\mathrm{SP_A}$	$\mathrm{T_{SA}}$ (s)	$\mathrm{SP_{SA}}$
NNC	13.000	8.00	33.683	3.088
BAC	6.333	16.42	27.347	3.803
HMC	8.533	12.19	25.657	4.053
MCS	16.067	6.47	24.941	4.170

ation of classifier systems in the case of semi-automatic classification processes. The specific application is for semi-automatic screening of the *Drosophila* heart. Computer vision techniques are employed to compute cardiac parameters such as the ratio between the diameters in end-diastole and end-systole, or the ratio between diastolic and systolic durations. Anatomic M-modes are used to reduce complexity by projecting the frames onto a line. A multiclassifier system is implemented to segment the M-mode image and compute the parameters. The system is described in detail and the three classifiers are discussed.

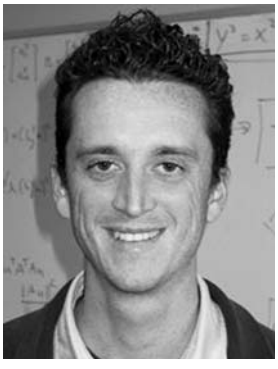
The resulting multiclassifier system shows better accuracy with respect to the single classifiers, but it requires more computational time. When the final objective is a completely correct identification of the end-diastolic and end-systolic points, the manual intervention of an expert user is required. As a consequence, the required time is compared between a complete manual segmentation and a semi-automatic solution in which the expert user is only asked to correct the points misclassified by the vision system. With the multiclassifier the time required to a medical expert for extracting the measure is decreased dramatically: from an average time of about 1 min and 40 s in the case of manual measurement, the procedure takes only about 25 s, including the time for correcting the errors of the automatic process. In such a manner the whole procedure is speeded-up and can be adopted in the current evaluation of the *Drosophila* heart's modifications at the Burnham Institute, La Jolla, California.

References

1. Paternostro G, Vignola C, Bartsch DU, Omens JH, McCulloch AD, Reed JC (2001) Age-associated cardiac dysfunction in *Drosophila melanogaster*. Circ Res 88:1053–1058
2. Lakatta EG (2001) Heart aging: a fly in the ointment? Circ Res 88:984–986
3. Frangi AF, Niessen WJ, Viergever MA (2001) Three-dimensional modeling for functional analysis of cardiac images, a review. IEEE Trans Med Imaging 20(1):2–25
4. McEachen JK II, Duncan JS (1997) Shape-based tracking of left ventricular wall motion. IEEE Trans Med Imaging 26(3):270–283
5. McVeigh ER (1996) MRI of myocardial function: motion tracking techniques. Magn Reson Imaging 14(2):137–150
6. Nahrendorf M, Hiller K-H, Hu K, Ertl G, Haase A, Bauer WR (2003) Cardiac magnetic resonance imaging in small animal models of human heart failure. Med Image Anal 7(3):369–375
7. Waiter GD, McKiddie FI, Redpath TW, Semple SIK, Trent RJ (1999) Determination of normal regional left ventricular function from cine-MR images using a semi-automated edge detection method. Magn Reson Imaging 17(1):99–107
8. Amini AA, Yasheng C, Elayyadi M, Radeva P (2001) Tag surface reconstruction and tracking of myocardial beads from SPAMM-MRI with parametric B-spline surfaces. IEEE Trans Med Imaging 20(2):94–103
9. Denney TS Jr (1999) Estimation and detection of myocardial tags in MR image without user-defined myocardial contours. IEEE Trans Med Imaging 18(4):330–344
10. Haber I, Metaxas DN, Axel L (2000) Three-dimensional motion reconstruction and analysis of the right ventricle using tagged MRI. Med Image Anal 4(4):335–355
11. Osman NF, McVeigh ER, Prince JL (2000) Imaging heart motion using harmonic phase MRI. IEEE Trans Med Imaging 19(3):186–202
12. Klein GJ, Huesman RH (2002) Four-dimensional processing of deformable cardiac PET data. Med Image Anal 6(1):29–46
13. Debreuve E, Barlaud M, Aubert G, Laurette I, Darcourt J (2001) Space–time segmentation using level set active contours applied to myocardial gated SPECT. IEEE Trans Med Imaging 20(7):643–659
14. Bosch JG, Mitchell SC, Lelieveldt BPF, Nijland F, Kamp O, Sonka M, Reiber JHC (2002) Automatic segmentation of echocardiographic sequences by active appearance motion models. IEEE Trans Med Imaging 21(11):1374–1383
15. Gerard O, Billon AC, Rouet JM, Jacob M, Fradkin M, Allouche C (2002) Efficient model-based quantification of left ventricular function in 3-D echocardiography. IEEE Trans Med Imaging 21(9):1059–1068
16. Jacob G, Noble JA, Behrenbruch C, Kelion AD, Banning AP (2002) A shape-space-based approach to tracking myocardial borders and quantifying regional left-ventricular function applied in echocardiography. IEEE Trans Med Imaging 21(3):226–238
17. Montagnat J, Sermesant M, Delingette H, Malandain G, Ayache N (2003) Anisotropic filtering for model-based segmentation of 4D cylindrical echocardiographic images, Pattern Recognit Lett 24(4–5):815–828
18. Clarysse P, Friboulet D, Magnin IE (1997) Tracking geometrical descriptors on 3-D deformable surfaces: application to the left-ventricular surface of the heart. IEEE Trans Med Imaging 16(4):392–404
19. Herment A, Mousseaux E, Dumée P, Decesare A (1998) Automatic detection of left ventricular borders on electron beam CT sequential cardiac images using an adaptive algorithm, Comput Med Imaging Graph 22(4):291–299
20. Kass M, Witkin A, Terzopolous D (1988) Snakes: active contour models. Int J Comput Vis 1:321–331
21. Staib LH, Duncan JS (1992) Boundary finding with parametrically deformable models. IEEE Trans Patt Anal Mach Intell 14(11):1061–1075
22. Chalana V, Linker DT, Haynor DR, Yongmin K (1996) A multiple active contour model for cardiac boundary detection on echocardiographic sequences. IEEE Trans Med Imaging 15(3):290–298
23. Yuille AL, Hallinan P (1992) Deformable templates. In: Blake A, Yuille AL (eds) Active vision. MIT press, Cambridge, pp 20–38
24. Cootes TF, Taylor CJ, Cooper DH, Graham P (1992) Training models of shape from sets of examples. In: Proceedings of British machine vision conference, pp 9–18
25. García-Fernández MA, Zamorano JL, Azevedo J (1997) Doppler tissue imaging. McGraw-Hill, New York
26. Malpica N, Santos A, Pérez E, García-Fernández MA, Desco M (2003) A snake model for anatomic M-mode tracking in echocardiography. In: Loncaric S, Neri A, Babic H (eds) Proceedings of 3rd international symposium on image and signal processing and analysis 2, pp 722–726

27. Mele D, Pedini I, Alboni P, Levine RA (1998) Anatomic M-mode: a new technique for quantitative assessment of left ventricular size and function. Am J Cardiol 81:82G–85G
28. Haralick RM, Shapiro LG (1985) Image segmentation techniques. Comp Vis Graph Image Process 29:100–132
29. Pal NR, Pal SK (1993) A review of image segmentation techniques. Pattern Recognit 26(9):1277–1294
30. Tobias OJ, Seara R (2002) Image segmentation by histogram thresholding using fuzzy sets. IEEE Trans Image Process 11(12):1457–1465
31. Manay S, Yezzi A (2003) Anti-geometric diffusion for adaptive thresholding and fast segmentation. IEEE Trans Image Process 12(11):1310–1323
32. Roli F, Giacinto G (2002) Design of multiple classifier systems. In: Bunke H, Kandel A (eds) Hybrid methods in pattern recognition. World Scientific Publishing, Singapore
33. Kittler J, Hater M, Duin RPW (1996) On combining classifiers. In: Proceedings of 13th international conference on pattern recognition, vol 2, pp 897–901
34. Woods K, Kegelmeyer WP Jr, Bowyer K (1997) Combination of multiple classifiers using local accuracy estimates. IEEE Trans Pattern Anal Mach Intell 19(4):405–410
35. Plessis B, Sicsu A, Heutte L, Menu E, Lecolinet E, Debon O, Moreau J-V (1993) A multi-classifier combination strategy for the recognition of handwritten cursive words. In: Proceedings of 2nd international conference on document analysis and recognition, pp 642–645
36. Dasarathy BV, McCullough CL (1998) Intelligent multi-classifier fusion for decision making in ballistic missile defense applications. In: Proceedings of IEEE conference on decision and control, vol 1, pp 233–238
37. Briem GJ, Benediktsson JA, Sveinsson JR (2001) Boosting, bagging, and consensus based classification of multisource remote sensing data. In: Josef Kittler, Fabio Roli (eds) Proceedings of 2nd international workshop on multiple classifier systems (LNCS 2096). Springer, Berlin Heidelberg New York, pp 279–288
38. Jain AK, Prabhakar S, Hong L (1999) A multichannel approach to fingerprint classification. IEEE Trans Pattern Anal Mach Intell 21(4):348–358
39. Chiou GI, Hwang J-N (1994) Image sequence classification using a neural network based active contour model and a hidden Markov model. In: Proceedings of international conference on image processing, vol 3, pp 926–930
40. Shah S, Sastry PS (2004) Fingerprint classification using a feedback-based line detector. IEEE Trans Syst Man Cybern B 34(1):85–94
41. Zhang X-P (2001) Thresholding neural network for adaptive noise reduction. IEEE Trans Neural Netw 12(3):567–584
42. Komo D, Chang C-I, Ko H (1994) Neural network technology for stock market index prediction. In: Proceedings of international symposium on speech, image processing and neural networks, vol 2, pp 543–546
43. Kuntimad G, Ranganath HS (1999) Perfect image segmentation using pulse coupled neural networks. IEEE Trans Neural Netw 10(3):591–598
44. Rae R, Ritter HJ (1998) Recognition of human head orientation based on artificial neural networks. IEEE Trans Neural Netw 9(2):257–265
45. Anthony M, Bartlett PL (1999) Neural network learning: theoretical foundation. Cambridge University Press, Cambridge
46. Lei T, Udupa JK (2003) Performance evaluation of finite normal mixture model-based image segmentation techniques. IEEE Trans Image Process 12(10):1153–1169
47. Acton ST, Mukherjee DP (2000) Scale space classification using area morphology. IEEE Trans Image Process 9(4):623–635
48. Giordana N, Pieczynski W (1997) Estimation of generalized multisensor hidden Markov chains and unsupervised image segmentation. IEEE Trans Pattern Anal Mach Intell 19(5):465–475
49. Pieczynski W (2003) Pairwise Markow chains. IEEE Trans Pattern Anal Mach Intell 25(5):634–639
50. Marroquin JL, Santana EA, Botello S (2003) Hidden Markov measure field models for image segmentation. IEEE Trans Pattern Anal Mach Intell 25(11):1380–1387
51. Cai J, Liu Z-Q (2001) Hidden Markov models with spectral features for 2D shape recognition. IEEE Trans Pattern Anal Mach Intell 23(12):1454–1458
52. Chien J-T (1999) Online hierarchical transformation of hidden Markov models for speech recognition. IEEE Trans Speech Audio Proc 7(6):656–667
53. Bao P, Zhang L (2003) Noise reduction for magnetic resonance images via adaptive multiscale products thresholding. IEEE Trans Med Imaging 22(9):1089–1099
54. Fjortoft R, Delignon Y, Pieczynski W, Sigelle M, Tupin F (2003) Unsupervised classification of radar images using hidden Markov chains and hidden Markov random fields. IEEE Trans Geos Rem Sensing 41(3):675–686
55. Ephraim Y, Merhav N (2002) Hidden Markov processes. IEEE Trans Inform Theory 48(6):1518–1569
56. Forney GD Jr (1973) The Viterbi algorithm. Proc IEEE 61:268–278
57. Viterbi AJ (1967) Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. IEEE Trans Inform Theory IT-13:260–269
58. Xu L, Krzyzak A, Suen CY (1992) Methods for combining multiple classifiers and their applications to handwriting recognition. IEEE Trans Syst Man Cybern 22(3):418–435
59. Huang YS, Suen CY (1995) A method of combining multiple experts for the recognition of unconstrained handwritten numerals. IEEE Trans Pattern Anal Mach Intell 17(1):90–94

Author Biographies



Luca Bertelli (S'04) received the D.Ing. degree cum laude in electronic engineering from the University of Modena, Italy in 2003 and the M.S. degree in Electrical and Computer Engineering from the University of California, Santa Barbara in 2005, where he is currently pursuing the Ph.D. degree in the field of Computer Vision and Pattern Recognition.His research interests include segmentation, perceptual grouping, level set methods and other aspects of machine learning.

Rita Cucchiara (M '89, PhD '92) is Professor at the Faculty of Engineering in Modena of the Università degli Studi di Modena e Reggio Emilia. She is the coordinator of the Ph.D. course in Computer Engineering and Science, of the Doctorate School in ICT (Information and Communication Technologies). Rita Cucchiara works on computer

vision systems, models and techniques for pattern recognition, image and video processing, and multimedia. Since 1998, she is responsible in Modena of the Imagelab research group (http://imagelab.ing.unimo.it). At Imagelab, she coordinates the research activity in video surveillance, people analysis, medical imaging for dermatology, and video analysis, segmentation and annotation for multimedia applications. These activities are carried out within the framework of many different national and international projects, funded by public agencies and private companies. Rita Cucchiara is author or co-author of 32 papers on international journals, 4 papers on national journals, 4 book chapters, and more than 100 papers on national or international conference proceedings. Since many years, she participates in the Program Committee of the main conferences on pattern recognition and computer vision (IEEE CVPR, IEEE ICME, ICPR, MIR, ACM Multimedia). She acts as reviewer of the main journals in the field.She is Associate Editor of Multimedia Tools and Applications journal and of Machine Vision and Applications journal. She is a member of GIRPR (Gruppo Italiano di Ricercatori in Pattern Recognition-associated IAPR), AI*IA , ACM and IEEE Computer Society. In 2006 she has been nominated IAPR fellow.

Giovanni Paternostro earned his Ph.D. in Biochemistry from the University of Oxford, England, in 1997. He has obtained his MD and Board Certification in Cardiology from the University of Rome, Italy. After postdoctoral training at the Imperial College School of Medicine, Hammersmith Hospital, London and at the Burnham Institute he was promoted to Research Investigator in 2001 and to Assistant Professor in 2003. Since 2005 he is also Adjunct Assistant Professor in the Department of Biomedical Engineering, UCSD. His research has been recognized by the 2002 Society for Geriatric Cardiology Basic Science Award and by the Ellison Medical Foundation New Scholar in Aging Award. In his research he uses a systems biology approach to investigate aging and heart disease in animal models and in humans.

Andrea Prati (M'98, PhD'02) received the Laurea degree (magna cum laude) in computer engineering in 1998 and the Ph.D. degree in computer engineering, both from the University of Modena and Reggio Emilia, Italy, in 2002. During the final year of his Ph.D. studies, he spent 6 months as visiting scholar at the Computer Vision and Robotics Research (CVRR) Lab at the University of California, San Diego (UCSD), working on a research project for traffic monitoring and management through computer vision.He is currently an Assistant Professor with the Faculty of Engineering, Dipartimento di Scienze e Metodi dell'Ingegneria, University of Modena and Reggio Emilia. His research interests are mainly on motion detection and analysis, shadow removal techniques, video transcoding and analysis, computer architecture for multimedia and high performance video servers, video-surveillance and domotics. He is author of more than 70 papers in international and national conference proceedings and leading journals and he serves as reviewer for many international journals in computer vision and computer architecture. Dr. Prati is a member of the IEEE, ACM and IAPR.