Biometric Recognition Based on Line Shape Descriptors*

Anton Cervantes[1], Gemma Sánchez[1], Josep Lladós[1], Agnès Borràs[1], and Ana Rodríguez[2]

[1] Centre de Visió per Computador i Departament de Ciències de la Computació
Universitat Autònoma de Barcelona, Edifici O, Campus UAB
08193, Bellaterra, Catalonia, Spain
{anton, gemma, josep, agnesba}@cvc.uab.es
[2] Hospital Universitari Arnau de Vilanova
Lleida, Catalonia, Spain
arodriguez@arnau.scs.es

Abstract. In this paper we propose biometric descriptors inspired by shape signatures traditionally used in graphics recognition approaches. In particular several methods based on line shape descriptors used to identify newborns from the biometric information of the ears are developed. The process steps are the following: image acquisition, ear segmentation, ear normalization, feature extraction and identification. Several shape signatures are defined from contour images. These are formulated in terms of zoning and contour crossings descriptors. Experimental results are presented to demonstrate the effectiveness of the used techniques.

1 Introduction

Biometric technology is based on identifying one individual from another by measuring some unique features like face, iris, voice, DNA, fingerprint or ear shape. A number of contributions exist in the literature presenting mature solutions using the above biometric descriptors. The reader is referred to [1] for a good introduction to biometric recognition. Although being apparently different research fields, biometrics and graphics recognition are in some cases close areas, at least from the methodological point of view. Some biometric descriptors, in particular fingerprints consist of line structures spatially arranged. A fingerprint structure encoded, and matched, in terms of the geometry and topology of ridges and minutiae is somehow equivalent to a line drawing consisting of lines and junctions. Probably due to these "close" representation, some authors have experimented with similar techniques in document analysis and biometrics. For example, Bunke and his team have applied graph matching techniques, often used in symbol recognition, in fingerprint classification [2]. Govindaraju et al. [3] used chaincodes, a typical representation in line drawings, for fingerprint matching. In addition to fingerprints, a biometric descriptor traditionally

* This work has been partially supported by the Spanish project CICYT TIC 2003-09291.

W. Liu and J. Lladós (Eds.): GREC 2005, LNCS 3926, pp. 346–357, 2006.

familiar among the Document Analysis community, other biometric descriptors can also be formulated from a "graphics recognition" perspective. Actually it is necessary to represent the biometric descriptor using geometric and structural properties of basic features as lines (contour approximation) or characteristic points. The recognition is therefore formulated in terms of shape similarity. Other examples are the shape of vessels [4] or the ear shape [5], [6]. Ear shape is another non-intrusive biometric descriptor that can also be formulated in terms of line-to-line matching. In particular, in this paper we propose the ear shape for the identification of newborns.

The use of biometric approaches for newborn presents several drawbacks. DNA is invasive so it can not be used each time the baby is changed of room. Iris pattern and retina are also invasive because the first days newborns have their eyes closed so taking images is very difficult. Foot geometry is not characteristic enough in the first days of life, and hand geometry is difficult to acquire because newborns usually have their hands closed and keeping all the fingers in the correct position is not easy. In our research we have tested two different approaches: fingerprints and ear shapes. Using fingerprints we could check that although fingerprints are fully formed at about seven months [7] of fetus development, the first days of life fingerprints seem not to be mature enough to be acquired properly. In the case of ear shapes, the two main reasons to choose this biometric descriptor are because it has enough recognition ratio for our purpose and because while other techniques historically associated to newborn identification (for example footprints) are not 100% passive, ear acquisition does not require any kind of cooperation. The results obtained working with fingerprints and ear shapes are in [8].

Ear shape is not as discriminant as other features but in some frameworks could be more suitable. In the literature there exist some approaches about the ear features extraction. In [5] a set of circles are created and centered in the centroid of the contours of the ear, and a count of the number of the intersection points for each radius and all the distances between neighboring points are used in the recognition process. In [6] a geometrical vector containing normalized distances between characteristic points of the inner ear is used, a vector describing the outer ear contour is also used to compare ears. In [9] each ear is modelled as an adjacency graph built from the Voronoi diagram of its curve segments and a graph matching process is performed. In [10] a linear transform that transforms an ear image into a smooth dome shaped surface whose special shape facilitates a new form of feature extraction that extracts the essential ear signature without the need for explicit ear extraction is developed. Our work proposes a set of shape signatures designed to describe ear shapes and to be used in a biometric newborn identification framework. Some interesting reviews on general shape recognition have been looked up and some of the descriptors in this paper proposed are inspired on techniques explained in [11], [12] and [13].

The general organization of our approach consists in the following steps: image acquisition, ear segmentation, ear normalization, feature extraction and identification process. In the ear segmentation step a preprocessing to enhance ear

Fig. 1. Grey-level images

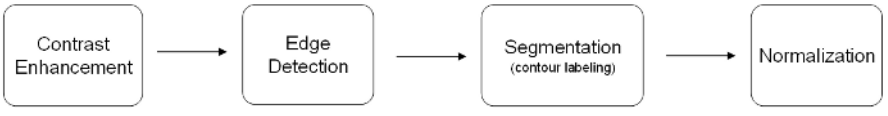


Fig. 2. A flowchart of the proposed algorithm

edges and a processing of the *Canny* edges using some typical characteristics of the ear to get the chain around the ear is performed. In the ear normalization step an affine transformation to standardize ears before the feature extraction process is performed. In the feature extraction process the distribution of some characteristic points is studied using zoning and contour-crossings based descriptors. Finally, in the identification step an algorithm to reach the most similar class in the database using different signatures is explained.

The organization of this paper is as follows: in section 2 the image acquisition, ear segmentation and the normalization process algorithms are developed, in section 3 the descriptors used to describe the ear structure are explained. In section 4 the recognition algorithm to reach the final decision is described, in section 5 the experimental results obtained with our own database are presented, finally in section 6 conclusions and future work are explained.

2 Ear Segmentation and Normalization

Ear images have been acquired with a high resolution digital camera. The model of the camera is a Nikon Coolpix 4300 and has a resolution of 4 megapixels. The images acquired have been cropped (512x512) to obtain images where only the ear appears, this way the following step, the ear segmentation process, becomes easier. Illumination conditions were controlled trying to avoid brightness points. Some examples of grey level images are shown in Fig 1.

After acquiring the image four different steps are performed before the feature extraction could be done, a flowchart of the proposed algorithm is shown in Fig 2. In the first one a preprocessing to improve the outer ear contour location is performed. In the second one the contours of the image are obtained using an edge detector. In the third step the edge segmentation is developed after

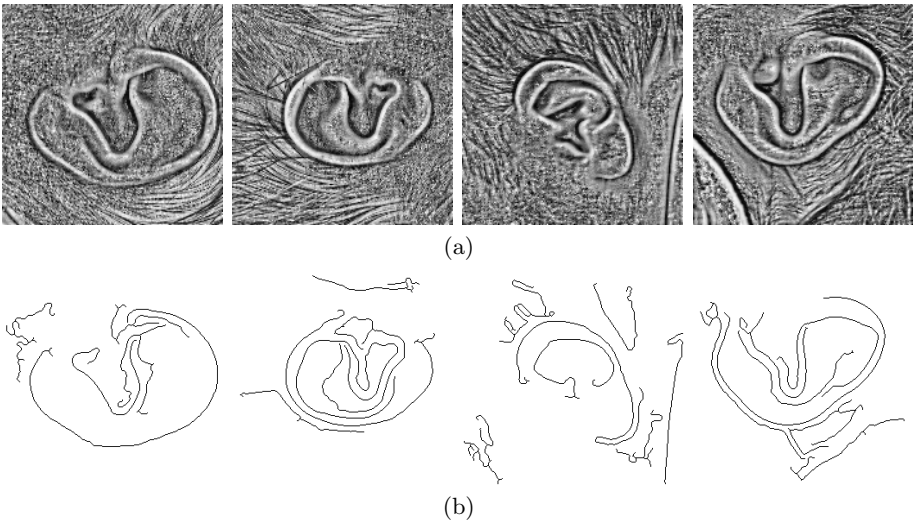


Fig. 3. (a) Local maximization of the contrast of the grey level images (b) Edges found by the *Canny* detector applied over the the images of a)

labelling and processing the pixels obtained in the previous step. Finally the ear normalization is computed. Let us further describe each step of the process:

The aim of the first part of the process consist of a little preprocessing of the grey level image to enhance the outer ear contour. For this purpose a local maximization of the contrast is applied before the edge extraction to obtain a image where the outer ear contour becomes more contrasted, see Fig 3 (a). Once this enhancement is performed getting the edge around the ear becomes easier. The window used in the enhancement process is square because at this moment of the process the position and orientation of the ear in the image is still unknown. The size of the window is empirically set.

In the second step of the process, after the local maximization of the contrast is computed the edge map is obtained using the *Canny* operator in the preprocessed image. The sigma value of the *Canny* detector should be great enough to remove spurious artifacts but small enough to avoid unnecessary smoothing to preserve the original shape and getting the correct location of the ear. In spite of the smoothing carried out by the edge detector many spurious lines as freckles, possible changes in color of the skin, hair and so on also appear in the line structure. Visualizing the edges obtained is easy to see that the line around the ear is one of the longest ones. Therefore a cleaning process removing lines whose length is lower than a fixed threshold is performed. After this process an image with the contours of the ear and spurious lines that have survived to the removing process is obtained, see Fig 3 (b).

In the third step of the process, the ear segmentation is performed after a pixel labelling carried out depending on some characteristics as number of neighbors and level and sign of curvature computed along the edges. In the labelling process

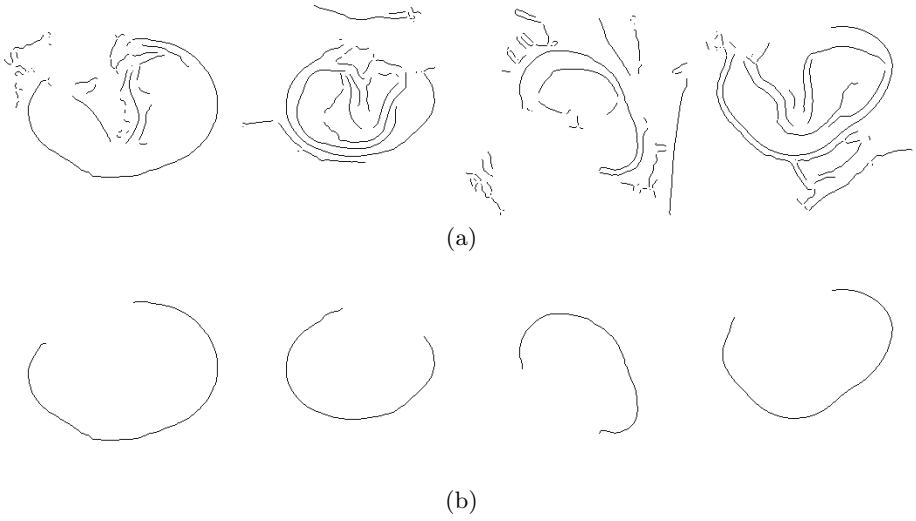


Fig. 4. (a) *Canny* image without the pixels whose curvature is higher in magnitude than a threshold. (b) Longest line (segmentation).

pixels are grouped into three different classes. The first class of pixels is named P_b and contain pixels that have more than two neighbors (bifurcation pixels), the second class is P_{hc} and the curvature level (magnitude) of these pixels is higher than a predefined threshold, *thr*. The rest of the pixels are the pixels whose curvature is lower than thr, and are assigned to class P_{lc}. This threshold is image dependent, so that the contours obtained with the *Canny* operator depends on how far the image has been taken. The pixel classes have been selected this way because in general, the outer ear contour does not contain pixels from P_b class nor from P_{hc}. For these reasons the algorithm steps to find the correct edge is as follows. First of all a removing step of the pixels of P_b class is performed. Next, the curvature along the edges remaining in the image is computed, and the pixels belonging to P_{hc} are also removed, see Fig 4 (a). After performing these two previous stages, only smooth shapes remain in the image, in Fig 5 (d) P_b, P_{hc} appear in black and P_{lc} are painted in grey. Finally, two images are extracted from the surviving pixels. The first one contains pixels whose curvature runs on $I_1 = [-thr, 1]$ and the second one contains the pixels whose curvature runs on $I_2 = [-1, thr]$. The first image contains pixels whose curvature is negative or with a positive value but with a magnitude under 1. In the second image it happens the contrary. This separation has been inspired by the morphology of the ear. The outer ear contour is similar to an ellipsoidal shape, therefore the sign of the curvature should remain unchanged almost everywhere. The outer ear contour is then assigned to the longest chain among all the edges of the two images, see Fig 4 (b). Two examples illustrating the segmentation reached whether the high curvature pixels are not removed or whether the pixels are not classified by the sign of their curvature are shown in Fig 5 and 6.

Fig. 5. (a) *Canny* edges (b) Longest edge before removing pixels of high curvature and pixels with more than two neighbors (c) Longest edge obtained using the information displayed in d), i.e., after removing pixels of high curvature and pixels with more than two neighbors (d) Black: bifurcation and high curvature pixels. Gray: Pixels with the magnitude of the curvature under a fixed threshold.

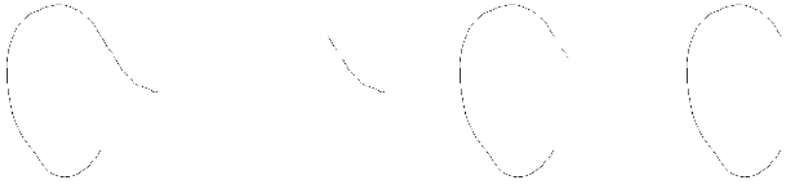


Fig. 6. (a) Segmentation without using the sign of the curvature (b) Pixels where the curvature is in I_1 (c) Pixels where the curvature is in I_2 (d) Longest edge among the edges from b) and c) (better approximation of the ear)

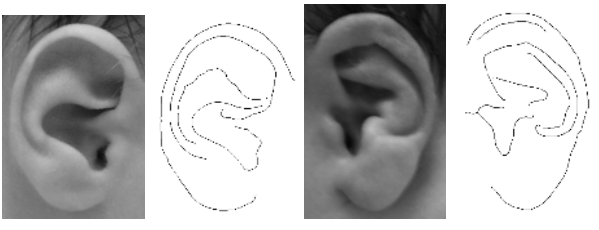


Fig. 7. (a) (c) Ear normalizations (b) (d) *Canny* map, the edges 'outside' the outer ear contour have been removed automatically

Finally, the normalization process is performed. This step consist of three different parts: a cropping, a rotation and a scaling. Cropping, to eliminate everything out of the ROI, rotation, to get the ear vertical, scaling to obtain similar sizes for all the images of the database. Some normalized images are shown in Fig 7 (a) and (c) and the *Canny* operator computed over the normalized images without short lines and lines out of the outer ear contour are in (b) and (d).

The curvature along all the lines of the image has been computed in each pixel as the difference between the mean orientation of the pixels before and after the current pixel divided by the size of the neighborhood used. The size of the neighborhood depends on what kind of points should be detected. Small neighborhoods must be used to find sudden changes of direction while greater neighborhoods must be used when smoother changes must be found. In our case a neighborhood of 10 pixels has been used.

3 Feature Extraction

The aim of this process consists in extracting some features to compare different ears in the recognition process. The features are extracted from the *Canny* edges computed over the normalized images. To obtain normalized features the major and minor axes are computed following the approach of [6]. These are used during the classification step as reference data to align the input image to the model. The major axe, Ay, is set as the segment joining the two furthest points of the outer ear, the center, O, is defined as the midpoint of Ay. The second axe, Ax, is the orthogonal to Ay by O, see Fig 8(a).

In this paper four different shape descriptors are used to characterize the ear contours:

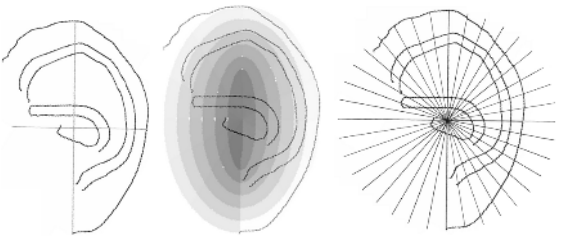
- Z_d density zoning features.
- Z_α angle zoning features.
- C contour-axes crossings
- E elongation (major-minor axes ratio).

Each descriptor is classified regarding to models and the corresponding outputs are combined to get the final decision. Let us further describe the feature descriptors defined to characterize a shape:

3.1 Zoning-Based Descriptors

Zoning is a well-known technique that describes a shape in terms of the spatial distribution of a set of feature points along a predefined lattice of regions or zones covering the shape. In our case, as ears usually have a shape similar to an ellipse, an ellipsoidal mask to extract the features of the same scale as the ear is constructed Fig 8(b). The ellipse is divided into 12 sectors, i.e, 30 degrees for each sector and each of these ones in 6 concentric rings. Two zoning descriptors are defined. Let us describe them.

1. The first feature extracted using the zoning mask is the density of edge points in each zone. This characteristic is computed dividing the number of pixels of the edge inside the zone by the number of pixels of the whole zone. The Euclidean distance is used to compare zonings of different ears in the recognition process:



(a) (b) (c)

Fig. 8. (a) Ear Axes (b) Zoning Mask (c) Contour-Axes Crossings

$$d_1(z_1, z_2) = \sum_{i=1}^{n} |z_1^i - z_2^i| \tag{1}$$

where z_j^i is the value of the i-zone of the j-zoning and n is the number of zones of the mask.

2. The second feature extracted using the zoning mask is the mean angle of the orientation of the edge pixels in each zone. The mean angle, α_m, is computed as follows:

$$\alpha_m(\alpha_1, ..., \alpha_n) = \frac{1}{2} \arctan \frac{\sum_{i=1}^{n} \sin(2\alpha_i)}{\sum_{i=1}^{n} \cos(2\alpha_i)} \tag{2}$$

Where $\alpha_i \ i = 1, ..., n$ are angles running on [0,180] and α_m is the mean orientation. The distance between two angles is given by:

$$d_\alpha(\alpha_1, \alpha_2) = min(|\alpha_1 - \alpha_2|, 180 - |\alpha_1 - \alpha_2|) \tag{3}$$

Where α_1 and α_2 are two angles running on [0,180]. The distance used to compare these zonings is:

$$d_2(z_1, z_2) = \sum_{i=1}^{n} d_\alpha(z_1^i, z_2^i) \tag{4}$$

3.2 Contour-Axes Crossings

It is another well-known shape signature. In general the idea is to measure the distribution of crossings between contour shapes and a set of reference axes strategically distributed. In our case 17 radial straight lines regularly distributed at a frequency of 10 degrees and centered at O. This descriptor computes the distribution of the intersections between ear diameters and the *Canny* edges in the inner ear, see Fig 8(c). To produce a normalized distribution the value assigned to each intersection is saved as its Euclidean distance to O divided by the half of the length of the major axe. Depending on the side where the intersection is found a different sign is assigned: if the intersection is over Ax the positive sign is preserved and if the intersection is under Ax the sign is changed

to negative. This way the values of the distribution are always on [-1,1]. For each ear a vector with the information about the intersections is saved in the database. The vector is the following :

$$V = \{[O_0, I_0^1, ..., I_0^{n0}], ..., [O_i, I_i^1, ..., I_i^{ni}], ..., [O_{17}, I_{17}^1, ..., I_{17}^{n17}]\} \quad (5)$$

Where O_i is the orientation of the i-diameter, I_i^j is the $j-$intersection in the i-diameter, n_i is the number of intersections in the i-diameter.

Given two contour-axes crossing signatures, the similarity between them is computed as follows. The Euclidean distance between all the intersections of one distribution is computed over all the intersections of the other distribution. Next, the smallest distance is taken and if this value is under a predefined threshold this distance is saved. The elements that produce this minimum value are removed (the elements have matched). This process is repeated until one of the distributions has no elements or until the minimum distance is higher than the predefined threshold. At this point, a vector with the distances of the matching intersections is obtained. The maximum value of this vector is selected as the distance between the distributions.

4 Identification Process

Using the descriptors explained in section 3 a several number of combinations of descriptors (signatures) have been tested to check which signature is the best to recognize newborns in our framework. The different signatures tested have been: $S_1 = [C]$, $S_2 = [Z_\alpha]$, $S_3 = [Z_d]$, $S_4 = [E]$, $S_{1,2} = [C, Z_\alpha]$, $S_{1,3} = [C, Z_d]$, $S_{2,3} = [Z_\alpha, Z_d]$, $S_{1,2,3} = [C, Z_\alpha, Z_d]$ and $S_{1,2,3,4} = [C, Z_\alpha, Z_d, E]$, see section 3 for an explanation of each descriptor.

To compare the values of the different components the distances and measures of similarity explained in the previous section are used: d_1 for density zonings, d_2 for angle zonings, the measure of similarity explained in section 3.2 for contour-axes crossings and finally the *Euclidean* distance is used to compare elongations.

Once all the descriptors of the ear models have been computed and stored in the database, the final result in the recognition process using a new ear image is obtained as follows (an example using the signature $S_{1,2,3,4} = [Z_d, Z_\alpha, C, E]$ with five models registered in the database is shown in Table 1). For each component of the current ear the corresponding similarity value with the same component of all the models in the database is obtained and with it a score is computed. The class whose similarity value is the lowest one is assigned the value 1, the second one is assigned the value 2 and so on (rows in Table 1). This process is performed for all the components of the signature. The scores each class of the database have obtained are summed (columns in Table 1). For each class of the database a similarity value is obtained. Now, a new similarity vector is obtained (last row in Table 1). The class with minimum sum of scores is returned. In the example of Table 1 the model returned by the algorithm would be the model *Ear3* which has the smallest sum of scores, 9.

Table 1. Score table

Current Ear	Classes inside the database Ear1	Ear2	Ear3	Ear4	Ear5
Z_α	2	3	1	4	5
Z_d	1	3	5	4	2
C	3	5	1	2	4
E	4	3	2	1	5
Sum of scores	10	14	9	11	16

5 Experimental Results

Although our project is oriented to newborn recognition the experiments for quantitative evaluation have been carried out using an adult ear database. This fact has been motivated because the project is in a preliminary stage and the newborn ear database is not great enough to obtain representative results. In spite of this fact the analysis performed over the newborn images that we have collected until now let us think that the recognition results obtained using adult images will be very similar or outperformed using the newborn ones due to the great variety of ear shapes observed in newborn ears.

As it has been said above the experiments have been carried out using our own database of adult ear images. In this moment the database consist of 140 images of 14 people. All the images were taken with the camera perpendicularly to the ear and avoiding brightness points. Trying to emulate a real process one of the images randomly selected of each ear is used as model in the registration process, and only one ear per person have been used in the recognition step. The rest of the images acquired have been used in the identification process.

The individual recognition ratios of each descriptor used and the ratio of the different signatures tested are shown in Table 2 and are graphically displayed in Fig 9. In both cases we can observe that results are grouped into three different groups: 1 NClass, 2 NClass, 3 NClass. A group i NClass shows the recognition values obtained using the i nearest classes, using the class ordination produced after computing the voting scheme among the input image and the registered ones.

Studying the obtained results the following conclusions can be extracted:

– Analyzing the recognition results obtained by S_1, S_2, S_3 and S_4 (signatures using only individual descriptors) it can be observed that S_1, S_2 and S_3

Table 2. Identification percentages of the whole set of signatures

	S_1	S_2	S_3	S_4	$S_{1,2}$	$S_{1,3}$	$S_{2,3}$	$S_{1,2,3}$	$S_{1,2,3,4}$
1 NClass	68	71	73	40	76	76	78	86	82
2 NClass	81	86	79	51	87	90	93	97	86
3 NClass	86	90	87	55	94	96	98	99	92

obtain similar rates around 70% using only the nearest class (1 NClass), around 83% using the two nearest classes (2 NClass), and around 88% using the three nearest classes (3 NClass). However, the fourth signature, S_4, does not obtain similar results, but much lower. Using the 1 NClass the recognition rate is only around 40% and using the 3 NClass the rate does not arrive to 60%. With these values we can conclude that the descriptor used in the fourth signature, S_4, is not characteristic enough, i.e., not useful to use in ear recognition.

- Analyzing the recognition results obtained by the signatures that combine 2 descriptors, $S_{1,2}$, $S_{1,3}$ and $S_{2,3}$, it can be observed that the recognition rates are also very similar among them and, in general, higher than the rates obtained by the signatures that only use one descriptor (S_1, S_2, S_3 and S_4). In this case the recognition values reached using the 3 NClass are between 94% and 98%.
- Finally, two signatures combining 3 and 4 descriptors have been tested. The signature that combines all the descriptors, $S_{1,2,3,4}$ uses the elongation which as it has been explained before it gets a very low recognition value. For this reason although being the signature that uses more descriptors is not the signature with the best results.
  The signature that combines the other three descriptors, $S_{1,2,3}$, obtains the best results and only using the 1 NClass the recognition ratio is higher than 85%. Using the 2 NClass and 3 NClass the recognition ratio is around 97% and 99% respectively.

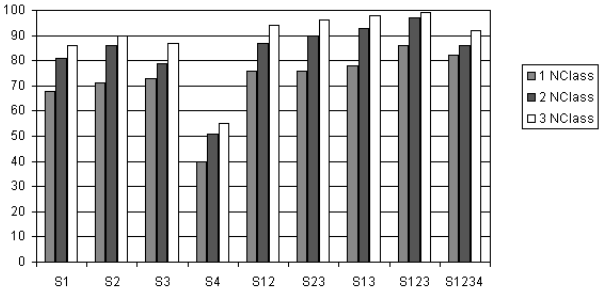


Fig. 9. Graphic of recognition percentages using the different signatures

6 Conclusions

In this paper an algorithm to recognize newborns using the biometric information extracted from the ear has been proposed. The algorithm presented begins with the image acquisition, goes on with the ear detection and normalization process of the ear an ends with the recognition process.

In the recognition process some combinations of different shape descriptors (signatures) have been tested to check which signature achieve the best results.

As can be observed in section 5 in general the signatures with only one descriptor achieve similar recognition ratios around 70% using only the nearest class. In the case of signatures that combine two descriptors it happens the same, this time around 77% using only the nearest class. The worst results are obtained using S_4 which is the signature that uses only the elongation of the ear in the recognition process, while the best results are obtained using $S_{1,2,3}$, where only using the 1 NClass the recognition ratio is higher than 85% and using the 2 NClass and 3 NClass the recognition ratio goes up to 97% and 99% respectively. Therefore, we conclude that a first approach to newborn recognition using the ear shape can be performed using these descriptors.

References

1. Anil K. Jain, 'An Introduction to Biometric Recognition'. IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Image- and Video-Based Biometrics, Vol. 14, No. 1, January 2004.
2. A. Serrau, G.L. Marcialis, H. Bunke and F. Roli. 'An Experimental Comparison of Fingerprint Classification Methods Using Graphs'. GbRPR. pp. 281-290, 2005.
3. V. Govindaraju, Z. Shi and J. Schneider. 'Feature Extraction Using a Chaincoded Contour Representation of Fingerprint Images'. AVBPA. pp. 268-275, 2003.
4. Z. Xu, X. Guo, X. Hu, X. Chen and Z. Wang. 'The Recognition Based on Shape for Blood Vessel of Ocular Fundus'. Proceedings of Sixth IAPR International Conference on Graphics Recognition (GREC2005), pp. 129-135, 2005.
5. M. Choras 'Ear Biometrics Based on Geometrical Method of Feature Extraction'. AMDO 2004, LNCS 3179, pp. 51-61,2004.
6. Z. Mu, 'Shape and Structural Feature Based Ear Recognition'. Sinobiometrics 2004, LNCS 3338, pp. 663-670, 2004.
7. W.J. Babler. 'Embryologic Development of Epidermial Ridges and Their Configuration'. Birth Defects Original Article Series, vol. 27, no. 2, 1991.
8. A.F. Cervantes 'Biometric Newborn Identification'. Master Thesis, Universitat Autnoma de Barcelona - Computer Vision Center, september, 2005.
9. M. Burge 'Ear Biometrics' Johannes Kepler University, Linz, Austria 1999.
10. David J. Hurley, 'Force Field Feature Extraction for Ear Biometrics'. Computer Vision and Image Understanding 98 (2005) pp. 491-512.
11. S. Loncaric 'A Survey of Shape Analysis Techniques' Pattern Recognition. Vol 31, No 8, pp. 983-1001, 1998.
12. D. Zhang 'Review of Shape Representation and Description Techniques' Pattern Recognition. Vol 37, pp. 1-19, 2004.
13. Remco C. Veltkamp and Michiel Hagedoorn. 'State-of-the-art in shape matching'. Technical Report UU-CS-1999-27, Utrecht University, the Netherlands, 1999