Face Localization using Geometric and Skin Characteristics

M. Saaidia¹, S. Lelandais² and M. Ramdani³ ¹Dpt de G énie-Electrique, Univ. de T ébessa, Alg érie ²IBISC FRE CNRS2873, Uni. EvryVal Essonne, France ³Dpt. d'Electronique, Univ.de Annaba, Alg érie msaaidia@mail.univ-tebessa.dz, s.lelandais@iut.univ-evry.fr, mes_ramdani@yahoo.com

Abstract

Geometric and skin face characteristics were used, in this paper, to localize face in images. In first stage, well known horizontal symmetric characteristic of face is exploited to localize the vertical symmetry axis of the face, then to delimit its vertical position in the image. In second stage, a neural network, trained to recognize pixel with skin characteristics, is used to completely localize the face by delimiting the region surrounding the vertical symmetry axis which colour characteristics agree with skin ones according to TSL colour space. At this stage, we explore three strategies to perform region delimitation. Finally, a quantitative measurement criterion is used to record localization quality. Experiments of the proposed method were carried out on images of XM2VTS database and also on a set of non-standard images.

Keywords: face localization, face detection, autocorrelation, neural networks, TSL

1. Introduction

Face localization or more generally face detection is being of great importance not only in research aspects but also in human daily life. This was possible mainly through the great technological developments essentially in information sciences where the machines are faster, more complex and more efficient. Thus, for face recognition (identity check), face expressions analysis or to take movements of face parts into account (gesture communication); localization of face in image or in video acquired by various peripherals (cameras, scanner, infra-red...) is necessary to the achievement of these operations. Several ways were explored by the researchers which give a great number of methods. Classification of all these methods depends on the parameter's classification. According to Hielmas and Low [1], two principal approaches can be distinguished; the global approach which consists in entirely seeking the face and the components approach which consists in finding the face through the localization and the regrouping of its components (eyes, nose...). For M-H. Yang et. al., [2] there are four major classes; knowledge-based methods, feature invariant approaches, template matching methods and appearance based methods. According to one or the other of these classifications, each developed method exploits one or more characteristics of face like colour, shape, movement etc, to perform face detection.

In this proposed work we exploit two principal characteristics of the face to perform its localization in the image. Vertical symmetry, of the face is used firstly to determine the principal vertical symmetry axis in the image which is probably the vertical symmetry face's axis. To delimit the vertical zone of the face, autocorrelation function is applied to pixels on both sides of this axis. In second stage, colour and appearance skin characteristics were used

to search the face region within the previously delimited zone. In this goal a neural network trained to recognize pixel with skin characteristics was applied to pixels within the delimited region. Three strategies were explored and compared to determine the best way to perform the scan of that region. Offline measurement of method performance was done according to a quantitative measurement criterion [3]. The proposed method was experienced on XM2VTS database and a set of non- standard images.

In Section 2 we explain the proposed localization method then, in Section 3, the technique to determine method performances will be presented. Next, Section 4 will contain the experiments and comparison results. Finally, in Section 5 we will conclude.

2. Face localization Method

Symmetry characteristic of the face is the most apparent geometric particularity. It was usually used in researchers' works for face's detection [4, 5] or for the detection of feature's face [6, 7]. The advantage of this characteristic lies in its universality (independent of the ethnicity of the person) and its robustness against the different conditions and also perturbations produced by image's acquisition process like illumination, presence or absence of structural elements (glasses, moustache, ...), presence or absence of facial expressions, image quality, neutral or complex background, etc., (Figure 1).



Figure 1. Images from Different Databases (Yale, JAFFE and XM2VTS) with Different Quality and Acquisition Conditions

However, due to the fact that it's not restrictive to the face, this characteristic is still largely insufficient to resolve face's detection problem. Researchers use it in second phase of the detection process to confirm the first phase in which a region was localized as candidate to be a face [8] or to extract the eyes position on a region which is considered as a face [9]. In our proposed method we exploit this characteristic first to determine the principal symmetrical axis of the image using the fact that if it contains a face, the image's axis symmetry will be the same as the face's one. Then, pixels on both sides of this axis are correlated to extract the vertical region which contains the face.

2.1. Image Principal Symmetrical Axis

The presence of a face in an image, affects directly its symmetrical properties; essentially when it contains only one. Based on this observation, we propose an efficient and fast

procedure to determine the principal symmetrical axis, of the treated image, using correlation function according to equation (1).

$$C(j) = Cor(Mat1(I, j), Mat2(I, j))$$
(1)

where:

$$\begin{cases}
M at 1(I,j) = Mat(I, j) \\
and \\
Mat 2(I, j) = Mat(I, 2 * j - p + 1) \\
with : I = 1 : M, j = 1, 2, 3, ..., (N / 2), p = 1 : j
\end{cases}$$

Mat is an (MxN) matrix containing treated image values; Cor, the well known correlation function and C the obtained curve representing the energy evolution of the compiled correlation between the matrices Mat 1and Mat2.

On Figure 2, we give an example of images representing some steps of the compilation of C according to equation (1).



Figure 2. Image's symmetrical axis compilation (a) original image from Yale database, (b)sample's of pair images representing Mat1 and Mat2 evolution, (c) Correlation energy

The curve of Figure 2.c represents the evolution of the correlation energy which indicates the level of similarity between the two matrices *Mat1* and *Mat2*. The maximum level indicates the greatest similarity and therefore the position of the symmetrical axes of the image which will be probably the principal symmetrical axis of the face.

Figure 3, gives examples of symmetrical axis determined on images of different databases.



Figure 3. Symmetrical Axis Determination

The examples above are chosen to demonstrate the robustness of the symmetry characteristic of the face for different situations of image capture, like illumination and background nature; face gender, face expression and face position. These images were taken from the well known databases Yale, JAFFE, ORL and XM2VTS and from a set of non standard images used here to demonstrate more generalization.

2.2. Vertical Face Zone

The face's symmetrical axis being found, the vertical delineation of the face will be easily compiled according to equation (2)

$$C_{\rm lr}(j) = Cor(C_l(I, j), C_r(I, j))$$
⁽²⁾

where:

$$\begin{cases}
C_{I}(\mathbf{I}, \mathbf{j}) = Mat(I, Ja - j) \\
and \\
C_{r}(I, j) = Mat(I, Ja + j) \\
with : I = 1 : M, \\
j = 1, 2, 3, \dots until \cdots Ja + j = N \cdots or \cdots Ja - j = 1
\end{cases}$$

Vertical delineation of the face is based on the compilation of the correlation function, *Clr*, between the columns on both sides of the symmetrical axis, determined here by index *Ja*, found in first stage. Figure 4 shows an example of vertical delineation of the face according to the compilation of *Clr*.



Figure 4. Example of vertical delimitation of the face, (a) original image from XM2VTS database, (b) facial symmetrical axis determination, (c) vertical delineation of the facial region

According to the C_{lr} 's curve evolution, we determine its minimum value which will indicate the position of the break-line between face and non-face regions.

To demonstrate performances obtained in this step we show in Figure 5 some examples from the databases cited earlier.

Faces vertical delineations, in Figure 5, are examples of results which will be obtained applying equation 2 to images with different appearances like illumination, background, gender, position and size of the face.



Figure 5. Examples of Face Vertical Region Delimitation

Results in fifth, seventh and eighth images demonstrate that in some cases the developed algorithm fails to perform a correct delineation. This will be taken in account in the next step where we will delimit the face zone.

2.3. Face Region Localization

In this phase we will explore three ways to determine the exact area of the face inside or closest outside areas to the vertical boundaries obtained in the first phase.

All proposed approaches are based on the use of a neural network unit preliminarily trained to recognize pixel that their colour characteristics agree with those of a skin's one according to the TLS colour space representation.

Equation 3 gives TSL colour space formulation:

$$S = \left[\frac{9}{5(r'^{2} + g'^{2})}\right]^{1/2}$$

$$T = \begin{cases} \arctan(r'/g')/2\pi + 1/4, \dots g' > 0 \\ \arctan(r'/g')/2\pi + 3/4, \dots g' < 0 \\ 1/2, \dots, g' = 0 \end{cases}$$

$$L = 0.299R + 0.587G + 0.114B$$
(3)

Where r' and g' are the normalized components of R and G in the RGB colour space.

Indeed, Skin colour has proven to be a useful and robust cue for face detection, localization and tracking [10]. Skin colour modelling problem was widely studied for computer graphics, video signal transmission and compression, etc. These studies have led to the emergence of several colour spaces like RGB, HSI, YCrCb, YUV, TSL, ...

TSL colour space is a transformation of the normalized RGB one. Compared to other colour spaces, TSL appears more adapted for skin modelling [11]. This modelling way was mainly used for face detection [12, 13, 14].

For the method presented in this paper we choose to save time compilation by limiting the use of the colour characteristics, for skin recognize, only on the pixels of the region preliminarily delimited in first stage. So, three strategies are proposed to perform face delimitation.

2.3.1 Rectangular Model for Face Zone:

This was the first strategy adopted to obtain a rectangular zone containing the face in the treated image. The applied algorithm can be resumed in two steps:

(i) First we apply the neural network unit to the pixels, one by one, of the symmetrical axis in the image (in the two directions: pixel 1 to pixel N then pixel N to pixel 1) until we obtain the first skin's pixel on both sides. These two first pixels will be used as the two horizontal limits of the face in the image (Figure 6).



Figure 6. Examples of Rectangular Face Delimitation

(ii) In second step, a set of neural network units will be applied along the horizontal and vertical limits of the rectangular zone previously obtained to determine whether these pixels are skin pixels or not. At each iteration, we decide if we have to enlarge or to reduce the rectangle's limits according to the ratio of skin pixels on the treated limit (Figure 7).



Figure 7. Rectangular Face Zone Delimitation on Examples of Figure 6

On the complete images of Figure 7, the vertical and horizontal red lines are the initial delimitation lines obtained at the first step (i) and the white ones are the final delimitation lines given performing second step (ii). Black lines are intermediate results obtained going toward the final result.

2.3.2 Elliptical Model for Face Zone:

General shape of the face is very similar to that of an ellipse and even in the case where it is different; the components of the face can be gathered by an elliptical form. Therefore, elliptical model was largely used in algorithms for face detection and recognition [15, 16] In our second strategy; we choose to exploit this characteristic to enhance the delimitation procedure of the face's zone. To do so, the delimitation procedure exploits the principal symmetrical axis which was found at the first phase. Skin colour characteristic is then used to lead the delimitation procedure according to two steps:

(i) This step is the same as that in the first strategy. Except that the two first pixels obtained on both sides of the symmetrical axis are used here to delimit the major axis of the ellipse which will be compiled to encompass the face or at least the main features of the face. This will be done according to equation 4.

$$\begin{cases} x = x_{c} \pm \sqrt{b^{2} - \frac{b^{2}}{a^{2}} * (y - y_{c})^{2}} \\ a \approx \frac{2}{3} * b \end{cases}$$
⁽⁴⁾

Where :

 $\begin{cases} y & \text{the ordinates of the point } (x, y) \\ (x_c, y_c) & \text{the focus coordinates} \\ a \text{ and } b & \text{respectively the minor and major axis} \end{cases}$

Having the major axis b, we can directly determine the set of ordinates (y_i) and the coordinates of the focus point (x_c, y_c) . The relation between the minor and major axis, given in the second term of equation 4, is a supposition made according to the characteristics of the human faces [17]. So we use equation 4 to compile abscises set (x_i) and therefore the ellipse.

On Figure 8, we give results of applying this algorithm to the same set of images given on figure 6. White portion line indicates the major axis. The red curve is the ellipse or potion of ellipse compiled according to Eq.4 with some variations on the value of a to obtain real values for (x_i) .



Figure 8. Elliptical Delimitation on Examples of Figure 6

The preliminary delimitation is not very effective. In several cases, the ellipse is very large compared to the face and therefore it includes areas that are not part of the face and sometimes important parts are lost due to too tight delimitation.

(ii) In this second step, we will try to adjust the resulting contour, previously obtained, to properly bond to the contour of the face. This will be done by applying a neural network unit on each point of the elliptical contour obtained previously to determine whether or not this pixel is a skin one. In the case of a non-skin pixel we push this contour point to go towards the skin pixels according to the side on which the point is positioned. If not, the contour point is pushed to go towards the limits of the skin zone (Figure 9).



Figure 9. Elliptical Face Zone Delimitation on Examples of Figure 6

2.3.3 Contour Tracking of Face Zone:

An alternative to the proposed models will be the tracking of the pixels on the borders of the face zone. This can be achieved by exploiting only the first skin pixel find on the principal symmetrical axis compiled in the first phase. This procedure will be implemented in only one step:

First, we find the first skin pixel by applying a neural network unit on the pixels of the principal symmetrical axis of the image. Then, a set of eight neural network units are applied on the pixels surrounding this first skin-pixel to determine skin-pixels candidate to be on the borders of the face zone. To achieve tracking procedure, we define a priority order on the 8 directions of movement along the diagram shown in Figure 10.



Figure 10. Priorities Tracking Directions

Pixel 0 is the current treated pixel and surrounded by the set of eight pixels *Pixel i. pi* defines the priority order of the direction. To avoid blocking in a local loop, we define a procedure that generates a random direction in the case of a stoppage of evolution.

Applying this strategy, to images of Figure 6, gives delimited faces shown on Figure 11.



Figure 11. Faces Zone Delimitation Obtained by Contour Tracking

3. Method Performance Measurement

To give an objective appreciation of results given by the studied methods, a proposed way to calculate the detection rate based on the relation between the number of pixels correctly and wrongly detected as pixels of the face and the number of face pixels in each treated image was presented in [3]. To do so, all testing database images were manually segmented in three regions like it is shown in Figure 12.



Figure 12. Examples of Regions Definition. Top: original image, Bottom: mask of Regions

First region (white one on the masks of Figure 12) contains the W pixels which represent the essential components of the face (brows, eyes, nose, mouth and surrounding pixels). The second region (grey one) contains pixels surrounding the first region and belonging to the face. The last region contains all the B pixels of the image which do not belong to the face. For the detection system, the first region is one which has to be contained imperatively in the resulting contour and the third one has to be imperatively discarded from it. The second region is optional and has no effect on the computed results. We define two types of measures; Good detection rate (Gdr) and Quality detection rate (Qdr).

$$Gdr = \frac{W_1}{W}.100 \text{ and } Qdr = \left(\frac{W_1}{W} - \frac{B_1}{A - B}\right)100$$
 (5)

where W_1 and B_1 are respectively the number of pixels *correctly* and *wrongly* detected as belonging to the face and A is the number of all pixels of the image.



Figure13. Top: original images. Bottom: complementary relation between Gdr and Qdr measures

On Figure 13.a we have the same bad Qdr (about 41%), with two different Gdr(54% and 70%). On figure 13.b it's the same situation for a good Qdr (about 83%) with two different Gdr values (94% and 100%). To finish, we give in figure 13.c an example of a face perfectly detected with Gdr and Qdr at 100%. Thus, to have a correct appreciation of recorded results, each one of Gdr and Qdr has to be computed. Best results are obtained when they are both closest to 100% with minimum difference between them. However, from a rate Qdr equal to 70% and a rate Gdr equal to 90%, we can consider that the quality of detection is very acceptable.

4. Experimental Results

In order to check the validity of the proposed method studied here, experimental studies were carried out on the *XM2VTS* images database. This extended database contains 4 recordings of 295 subjects taken over a period of 4 months with rotating head shot in vertical and horizontal directions. Images are coloured and in ppm format. To determine principal symmetrical axes and vertical face zone we brought some transformations to original images like change to GIF format (more compressed) and the use of luminance information only (grey scale images). Instead, coloured images were used for other steps.

4.1. General Results

In this section we will give, discuss and compare the results obtained through experiments made according to the three strategies described above.



Figure 14. Comparison between Gdr measures for Rectangular (first curve), Elliptical (second curve) and contour tracking (third curve) delimitation procedures applied on 259 of XM2VTS images

The three curves given in Figure 14 were recorded for Gdr measure applied to the 259 first images of XM2VTS database according to Eq. 5. The average rates are also given. These curves show that the three strategies are valid for almost all images processed. Indeed, only 17 images for the first strategy, 16 for the second and 18 for third one have rates below 80% which represents about 6% of fail responses. Comparing the three curves, we can see that in the case of elliptical delimitation results are more regular and closer to 100%.

For the Qdr curves shown on Figure 15, the same remarks can be done. For the three curves, we have only 19, 21 and 19 images which have a Qdr below than 70%. This, represents a failure rate less than 9%.

Combined results between Gdr and Qdr measures show that the three strategies give a success rate (Gdr>=80% and Qdr>=70%) greater than 90%. The best result is given by the second strategy with about 92% of success.



Figure 15. Comparison between Qdr measures for Rectangular (first curve), Elliptical (second curve) and contour tracking (third curve) delimitation procedures applied on 259 of XM2VTS images

For the first strategy, success of the algorithm is based on the success of the initial vertical delimitation face zone but especially on the success of the first step (horizontal limits). For the second one, success is based on correct determination of the major axis of the ellipse. For the third one, it is tied to the success of pursuit procedure.

On Figure 16, we represent the time taken to process each image according to the three strategies. This time is only given for comparison between the three strategies.

Results show that the processing time is different for each image. This is due, firstly to the size of the face in the image, and secondly to the quality of the initial detection (rectangular or elliptical) of the face area for the two first strategies.



Figure 16. Comparison between Processing times (Pt) for Rectangular (first curve), Elliptical (second curve) and contour tracking (third curve) delimitation procedures applied on 259 of XM2VTS images

The first strategy gives the lowest Pt average (about 10.39 ms). The second performance is obtained by the third strategy which has also the best regular distribution along the database. To clarify the comparison between these three cases we give, in Figure 17, a sample of measures Gdr, Qdr and Pt obtained for the 48 first images of the XM2VTS database.



Figure 17. Comparison between Gdr (first plot), Qdr (second plot) and Pt (third plot) on the 48 first images of XM2VTS database

4.2. Compared Results

Results obtained for the three versions of the proposed method are compared to those reported, on the same database (XM2VTS) and with the same measurement criterion, in the work presented in [3]; were Zernike geometrical moments and Eigen-faces are used to train neural network.

| | Gdr mean % | Qdr mean % | Processing time |
|------------------------------|---------------|---------------|-----------------|
| Rectangular- delimitation | 94.39 | 82.05 | 10.39 |
| Elliptical-delimitation | 94.71 | 83.74 | 20.51 |
| Contour-delimitation | 92.52 | 83.05 | 15.08 |

Table1 resumes the mean results obtained for the three strategies of the proposed method. Elliptical-delimitation strategy appears to be the best in terms of Gdr and Qdr measures. However, rectangular-delimitation strategy appears to be the fastest.

| | Gdr mean % | Qdr mean % |
|-------------------------|---------------|---------------|
| Elliptical-delimitation | 94.71 | 83.74 |
| Zernike moments | 93.82 | 82.07 |
| Eigen- faces | 90.16 | 79.82 |

Table2. Performance Comparison: Proposed Method with those of Paper [3]

Table 2 demonstrate that results obtained in this work and those obtained in the precedent referenced work are very similar in the sense of "quality of detection". However, in the case of the present method, procedure of detection is simpler and promise to be faster. Indeed, the first stage is compiled using the well known correlation function and the second one is based on a trained neural network dedicated to skin pixel recognition according to TSL colour model. The referenced work uses Zernike geometric moments, which time compilation cost is higher than the one required for TSL parameters. This will produce more complex and slower neural networks.

4.3. Out of XM2VTS

In this section we give some examples of applying the proposed method on a set of non-standard images.



Figure 18. Sample of Results Obtained by Applying Proposed Algorithm to a Set of Non-standard Images

Figure 18 gives four images from a set of non-standard images which we use to test developed algorithms. The most important characteristics which differ from the XM2VTS standard database images are the non-neutral background, the large variability in face's dimensions and in face's position in the image.

5. Conclusion

A fast and efficient face localization method using geometric and skin characteristics was presented in this paper. Firstly we search a vertical symmetry on the image to determine the principal symmetrical axis of the face then a neural network trained to recognize skin pixels is used to delimit the face zone. To do so, three strategies were studied; rectangular, elliptical and contour face's delimitation models.

Recorded results and their comparison demonstrate that elliptical face's model delimitation gives best results in terms of quantitative measurement adopted. Processing time, also compiled, demonstrate that rectangular face's delimitation is faster but contour one is more regular along the database. We demonstrate that "Quality detection" results are similar to those obtained with a referenced method applied on the same database and according to the same measurement criterion. However, simplicity of the presented method increasingly promising to become very efficient when implemented on a hard support which will permit parallel processing.

References

- [1] E. Hjelmas and B. K. Low, "Face detection: A survey", Computer Vision and Image Understanding, vol. 83, no. 3, (2001), pp. 236-274.
- [2] M. -H. Yang, D. J. Kriegman and N. Ahuja, "Detecting faces in images: a survey", IEEE transaction on pattern analysis and machine intelligence, vol. 24, no. 1, (2002) January.
- [3] M. Saaidia, A. Chaari, S. Lelandais, V. Vigneron and M. Bedda, "Face localization by neural networks trained with Zernike moments and Eigenfaces feature vectors: A comparison", AVSS2007, (2007), pp. 377-382.
- [4] Q. B. Sun, W. M. Huang and J. K. Wu, "Face detection based on colour and local symmetry information", Third IEEE Int. Conf. on Digital Object Identifier, vol. 2, Issue 1, (1998), pp. 130-135.
- [5] J. S. Hyun, H. K. Mi, S. C. Yun and C. K. Nam, "Face detection using sketch operators and vertical symmetry", Int. Conf. on Flexible Query Answering Systems, Milan, vol. 4027, (2006), pp. 541-551.
- [6] S. H. Yea, Y. C. Hao, F. C. Po and Y. T. Cheng, "Face Detection with High Precision Based on Radial-Symmetry Transform and Eye-Pair Checking", AVSS'06, Sydney, (2006), pp. 62.
- [7] A. Hamzah, A. Fauzan and M. S. Noraisyah, "Face localization for facial features extraction using a symmetrical filter and linear Hough transform", Artificial Life and Robotics, Springer Japan, vol. 12, no. 1-2, (2008) March.
- [8] A. Hadid, M. Pietikänen and B. Martinkauppi, "Color-based face detection using skin locus model and hierarchical filtering", 16th Int. Conf. on Pattern Recognition, Quebec City, Canada, (2002) August, pp. 196-200.
- [9] W. Qiong, Y. Jingyu and Y. Wankou, "Face Detection using Rectangle Features and SVM", Int. Journal of Intelligent Systems and Technologies, (2006) Summer.
- [10] V. Vezhnevets, V. Sazonov and A. Andreeva, "A Survey on Pixel-Based Skin Color Detection Techniques", Proc. Graphicon, (2003).
- [11] J. C. Terrillon, M. N. Shirazi, H. Fukamachi and Akamatsu, "Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images", Int. Conf. on Face and Gesture Recognition, (2000), pp. 54–61.
- [12] J. Kovac, P. Peer and F. Solina, "Human skin color clustering for face detection", EUROCON 2003, Computer as a Tool, The IEEE Region 8, vol. 2, (2003) September, pp. 144-148.
- [13] J. P. Guerfi, Gambotto and S. Lelandais, "Ligne de Partage des Eaux pour l'extraction de visage dans l'espace de couleurs TLS", CORESA, (2005) November.
- [14] D. Brown, I. Craw and J. Lewthwait, "A SOM based approach to skin detection with application in real time systems", British Machine Vision Conference, (2001).
- [15] N. K. Kim, I. Choi and S. I. Chien, "Face detection using scan-line based Hough transform and MLP", IAPR workshop on machine vision applications, Nara, JAPON, (2002), pp. 314-317.

- [16] A. X. Han, R. S. Ma and Y. Li, "Face detection and recognition with SURF for human-robot interaction", IEEE Int. Conf. on Automation and Logistics, ICAL '092009, (2009) September, pp. 1946-1951.
- [17] M. Paula, R. Leonardo, J. Luciano and B. Maria, "Facial dimensions, bite force and masticatory muscle thickness in preschool children with functional posterior crossbite", Brazilian Oral Research, vol. 22, no. 1, (2008) January/March.
- [18] M. Saaidia, S. Lelandais, V. Vigneron and M. Bedda, "Face Detection by Neural Network Trained with Zernike moments", Proc. the 6th WSEAS International Conference on Signal Processing, Robotics and Automation, (2007), pp. 36-41.