EXPLOITING VORONOI DIAGRAM PROPERTIES IN FACE SEGMENTATION AND FEATURE EXTRACTION

Abbas Cheddad [1,1], Dzulkifli Mohamad [2] and Azizah Abd Manaf [3]

[1]*School of Computing and Intelligent Systems*
Faculty of Computing and Engineering
University of Ulster, BT48 7JL.
Northern Ireland. United Kingdom
cheddad@gmail.com

[2]*Faculty of Computer Science and Information System*
University of Technology Malaysia (UTM)
Johor. Malaysia
dzulkifli@utm.my

[3]*Faculty of Computer Science*
Malaysian Military Academy(ATMA)
Kem Sungai Besi,
Kuala Lumpur. Malaysia
azizah@atma.gov.my

Abstract

Segmentation of human faces from still images is a research field of rapidly increasing interest. Although the field encounters several challenges, this paper seeks to present a novel face segmentation and facial feature extraction algorithm for gray intensity images (each containing a single face object). Face location and extraction must first be performed to obtain the approximate, if not exact, representation of a given face in an image. The proposed approach is based on the *Voronoi Diagram (VD)*, a well-known technique in computational geometry, which generates clusters of intensity values using information from the vertices of the external boundary of *Delaunay triangulation (DT)*. In this way, it is possible to produce segmented image regions. A greedy search algorithm looks for a particular face candidate by focusing its action in elliptical-like regions. VD is presently employed in many fields, but researchers primarily focus on its use in skeletonization and for generating Euclidean distances; this work exploits the triangulations (i.e., Delaunay) generated by the Voronoi Diagram for use in this field. A *Distance Transformation* is applied to segment face features. We used the BioID face database to test our algorithm. We obtained promising results: 95.14% of faces were correctly segmented; 90.2% of eyes were detected and a 98.03% detection rate was obtained for mouth and nose.

Keywords: Biometric; Voronoi Diagram; Delaunay Triangulations; Distance Transformation; Features Extraction.

1. Introduction

Computer science is becoming an important central discipline for a variety of scientific fields, leading it to become an increasingly multidisciplinary research area. The capability of taking advantage of advances in computer science to model more *"human–like"* behaviour has granted added value to the field and thereby introduced new research areas. However, the *Human Visual System (HVS)* remains the most complete and efficient-yet-complicated vision system that has ever been seen. In fact, researchers in the area of computer vision are trying to emulate the essential functionalities of the HVS. The fact that a face cannot be described in words has sparked the conceptualization of new technology based on *machine vision*. *Face Recognition Technology (FRT)*, an example of that effort, is the study of the process of identifying a face in a given image pattern, segmenting it from the rest of the image (background), and then proceeding to a recognition phase. Indeed, the fact that faces can vary in terms of position, size, skin intensity values, lighting effects, facial expressions, the presence or absence of hair or glasses, and occlusion (faces with hidden part(s)) makes this a challenging exercise. All of these factors contribute to our inability to easily predict the appearances of faces.

Unlike fingerprints, a process such as face recognition could - at least in principle - be used to recognize people "passively," that is, without their knowledge or cooperation. Reference [1] is a very interesting article which discusses face recognition from an ethical perspective.

[1] Corresponding author Tel: + (44) 7907570340. E-mail address: cheddad@gmail.com

The success of a recognition technique depends to a certain extent on two things: the first is minimization of the computational burden and the second is maximization of the speed of the process that leads to an accurate result. One method used by some face recognition researchers is to manually crop the face area from every relevant image [2]. For practical applications, however, this method would not be helpful.

The face recognition process usually requires three steps [3, 4]. The first step involves locating the face area, a process known as *"face localization."* The second step involves extracting facial features; this step is very critical because the next and final step depends solely on its outcome. The final step is classification of the face image based on the features vector. However, the first and second steps may overlap, and in certain cases face features are extracted first in order to locate the face region. A process in which a face is first detected by a coarse location of facial features and then fine features are subsequently extracted is known as a bottom-up process [5].

A number of algorithms have been proposed to extract facial features based on different approaches. Reference [6] used a specific 3*3 mask for edge detection to crop faces from the background. The author's mask operations on the image utilize a closed contour. Images were all shot in a black background and a black cloth was draped around the subject's neck. Therefore, in this case, clipping the faces was a trivial task, and this method is considered one of the more ancient, direct techniques for face segmentation.

References [7, 8] used snakes or active contours, and represent a popular method of object boundary detection. In this work, a controlled continuity spline function is used to transform the shape of the curve so as to minimize the energy function relative to the initial state of the curve (energy minimization). The final curve will then mimic the external boundary of the object. The main difficulties in this method are the computational burden, which is quite high, the sensitivity to noise, the fact that a good initial point is very hard to estimate and that the method will always converge to a solution, whether or not this solution is the desired one or not.

Colour (RGB) transformation is another popular method which was recently invoked. It is a very fast algorithm. In Reference [2], an RGB colour matrix was nonlinearly transformed to YC_bC_r (Intensity, Chromatic blue and Chromatic red) space. Reference [9] followed a similar procedure, while Reference [10] chose the following system to convert from (RGB) to (Y, C_b, C_r):

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} (0.299)(0.587)(0.114) \\ (-0.169)(-0.331)(0.500) \\ (0.500)(-0.419)(-0.081) \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

Their algorithm resulted in 85.92% correct face detection. Reference [11] describes a method for extracting the skin area from an image using normalized colour information. Once the flesh region is extracted, its colour distribution is compared with a manually cropped and constructed model. Reference [12] obtained a 90% correct detection rate using an identical method. Reference [13] created a special hardware system; called the Triple-band System, which included a Near-IR (Infra Red) illumination generator. Their idea is based on the premise human skin has a special reflection in the near-IR illumination. In other works (e.g., [14]) face segmentation has been done using another colour space transformation, namely HSV (Hue, Saturation and Value) and shape information. First, skin-like regions are segmented based on the Hue and Saturation component information, and then a fine search in each of the regions is performed to detect elliptical shapes. This method normally gives large false alarms, which requires further processing either by considering other information or by making additional use of other techniques. Moreover, generating a skin colour model is not a trivial task. For example Reference [15] used 43 million skin pixels from 900 images to train the skin-colour model and Reference [16] manually segmented a set of images containing skin regions to generate a skin model.

Template matching technique was among the first pioneered algorithms. It was refined to develop a *deformable template matching* which was implemented to compensate for some of the drawbacks of former methods [17, 18]. In their paper [17], Wang and Tan used six templates. Two eye templates and one mouth template were used to verify a face and locate its main features; afterwards, two cheek templates and one chin template were employed to extract the face contour. The performance claimed to be favourable; however, it could not detect faces with shadow, rotation or bad lighting conditions. One thing that they presented as an advantage was that they were able to choose a relatively big step in feature matching so as to reduce the computation cost. While it holds true that the said *"big step"* reduces the computation burden, it can, however, result in the loss of other significant information. The interesting point here is that

the authors admitted this indirectly in their conclusion: "*We suspect that one reason for this (failure) is that our template does not include enough information to distinguish faces in very complex background*". Most importantly the algorithm gave false positives when rotations occurred, and when there were unwanted objects whose shapes were similar to ellipses. Creating a good template model is not an easy task, however, because it does not (in fact it cannot) take into account the variable appearance of faces. It suffers most when dealing with unknown face size and rotation. Resizing the model and rotating it moreover increases the processing time.

The first advanced neural network approach to report results on a large complex dataset was presented in [19]. Their system incorporated face knowledge in a "retinally" connected neural network. The neural network was designed to look at windows of 20*20 pixels (i.e. 400 input units). This group introduced a hidden layer with 26 units: 4 units to look at 10*10 pixel sub-regions, 16 to look at 5*5 sub-regions, and 6 look at 20*5 pixel overlapping horizontal stripes. The input window was pre-processed for lighting correction (a best fit linear function was subtracted) and histogram equalization. Neural network methods require a lot of training samples in order to increase their efficiency levels. The complexity of the network is considered to be a disadvantage because one cannot know whether the network has "cheated" or not. It is almost impossible to find out how the network determines its answers. Hence, this is also known as a black box model.

Reference [20] uses the Genetic Algorithm (GA) to select a pair of eyes from among various possible blocks. The fitness value for each candidate (face) is calculated by projecting it onto the eigenfaces space. Even though GA is accurate, like the "snakes" process it suffers from being a time-consuming technique, because it fires its chromosomes to each and every location.

More relevant to our proposed work is Reference [5], which suggests a symmetry-based method for face boundary extraction from a binarized facial image. Basically, the work details construction of Delaunay Triangulations (which are the dual of the Voronoi diagram) from points of an edged image. The property of each triangle was examined geometrically, and the so-called *J-Triangle* (Junction triangle) was identified. These types of triangles act as linkers to repair the broken edges. The purpose of this is to prevent the face boundary from being merged with the background. This method may experience a heavy computational load if the image size increases; it will also be extremely difficult to deal with images with a complex background. The proposed method has limitations in cases of rotation or if the face of interest is wearing glasses, as stated in the paper. It is also sensitive to noise and the presence of a beard. The demonstrated success rate for detection of facial features upon segmentation is 89%. In contrast, as we will show later, our proposed method employs the Voronoi Diagram and its properties on a select few unique feature points derived from the image histogram rather than on points associated with edges in the special domain. Therefore, the computational process for our method is much less than that of the method proposed in [5].

For an additional detailed survey of the different techniques, we direct the reader to the literature [21], as well as to [22], in which the recognition phase is discussed, for a more in-depth study.

1.1. *Image content segmentation*

Reference [23] describes an interesting survey on colour image segmentation and presents tables comparing the advantages and disadvantages of the major techniques. However, this section also warrants a brief explanation of image segmentation techniques.

Broadly speaking image segmentation techniques can be classified into two categories: namely boundary-based techniques and region-based techniques. It is interesting to note that most segmentation techniques involving gray scale images can be extended to colour images. The basic and most traditional method of segmentation is based on dynamic thresholding of the image histogram. The threshold value is determined by examining the valley falling in between two peaks in the histogram. A repetitive thresholding process yields segmented regions. Obviously this method is feasible only when the examined image generates a bi-model histogram. Various threshold points, for the same image, can be obtained by applying this method on different colour spaces resulting from RGB colour transformations. Region-based algorithms include region growing, region splitting and region merging. Each of these methods triggers a seed to analyze the pixel values. The region growing method starts with a pixel seed and then examines the neighbour pixels to determine any homogeneity in terms of brightness. The selection of the seed and the order of the scanning process are crucial to its performance. At the sight of region splitting method, we find that the whole image is taken as the initial fired seed; this image is then

split into four blocks covering homogeneous areas. Again these four blocks are considered as the new seeds and the operation is repeated until homogeneous regions are found. In this algorithm, the final segmentation tends to retain the shape of the data structure of the splitting mechanism. Region merging can be any one of the former two, but the difference is that homogeneous regions are merged together while the process is running. The nearest neighbour algorithm to K-means is a machine-learning algorithm that implicates voting by a pixel's neighbours for homogeneity. Edge detection is another means of image content segmentation. Different operators for this have been introduced, including Sobel, Prewitt, Laplacian, and Canny operators. All such methods share the goal of targeting abrupt changes in image intensities, but differ in the weights assigned to each value in the mask.

Reference [24] presented a practical implementation of the Voronoi Diagram (VD) to segment images based on selected feature points. These points are exclusively identified and selected after careful analysis of unique pixels residing along the image edges of high gradient magnitude. For the sake of experimentation, all pixels with a gradient magnitude larger than 75% were chosen. This process was followed by a normalized cross-correlation in which each window was centred with one of the selected points. If any window survived the matching, meaning that it did not have another intensity similar window, its point would be selected as a unique feature point. The authors claimed their proposal could be used in recovering distorted images and also in watermarking.

Reference [25], in an attempt to produce a new class of automatic passive type of biometrics, suggested the ear as the basis for a supplementary source of evidence in identification and recognition systems. Development of their ear biometric graph model proceeded through different stages. First, a deformable contour method was used to locate the ear area. This process was followed by applying a canny edge detector. For selective curve extraction, the edge relaxation method was applied, and only curves with a length greater than 10 pixels were kept. A generalized Voronoi diagram was next initiated to yield a neighbour graph that uniquely identifies the studied ear.

2. The Face Databases

Our experiments were conducted on a number of databases. Two of them are internationally-recognized face databases used for Biometric algorithm testing. The purpose of using different sets is to see the performance of the algorithm in an uncontrolled environment. For the sake of a fair comparison with other published works, we constrained our testing reported in this work on images from set 3.

- Set1: The database of the Olivetti Research Laboratory in Cambridge. This database includes 400 facial images corresponding to 40 different persons. Each image has the size of 92*112 with 256 gray-scale levels. Images were taken under different lighting conditions and different poses (See Fig. 21 in the Appendix).
- Set2: Our own face database is being created for this purpose using a CCD monochrome camera. The images shown in the Appendix are internal to our group (taken from this Laboratory).
- Set3: The BioID Face Database. This dataset consists of 1521 gray-scale images, each with a resolution of 384x286 pixels. Tables reported on this work showing detection rates are generated using this database. The BioID can be downloaded from the following link: http://www.humanscan.de/support/downloads/facedb.php
- Set4: Still images obtained from the World Wide Web (WWW). Some samples are shown in the Appendix.

Our experiments were carried out using PC Pentium 4 with a 2.80 GHz CPU and 256 MB of RAM memory. The algorithm was developed using MATLAB (v.6.5). Randomly selected samples were chosen (more than 300 images) using the computer from Set 3 (1521 images) and with which the detection rates reported in this work were generated. Certain tests were conducted on a simple background while others showed a complex background. Since posing was not restricted, a candidate could possibly take any position. The algorithm is also invariant to scale since it deals with regions rather than size or area, and with rotation, since it uses a distance measurement taken from between face features to orient the ellipse model accordingly, as shown in Fig.14 in the Appendix. Our algorithm runs in a very reasonable time, despite MATLAB being an interpreting language.

3. Face Detection Algorithm

Many papers have been published that perform face localization in colour images [16, 10, 13, 26, 14, 24, 15, 2, 11, 9]. All of these exploited the HSV (Hue, Saturation and Value) colour space by using the RGB to HSV transformation.

Few colour databases of faces are available online in comparison with to gray scale ones; therefore, we directed our attention to the latter. Unlike colour images, gray scale databases still face a huge challenge in detecting facial region(s) from of scenes with complex or even moderately complex backgrounds.

Our approach starts with two facts in mind that characterize face regions, which are:

- A face is approximated best by an elliptical shape. This will help in discriminating face regions.
- A face is characterized –generally– by a ramp transition of gray intensities. This phenomenon assists us in obtaining a closed contour that bounds the face boundary.

3.1. *System overview*

Our system's graphical user interface (GUI) is shown with all of its functionalities in Figure 1. We call it DFD (Delaunay Face Detection). The functions underlying the proposed system benefit from the Voronoi properties to detect and segment faces and their features. A block diagram of the system is shown in the Appendix (Fig. 23). The first step in every machine vision task is the pre-processing phase, which includes but is and not limited to: *Histogram Equalization*, *Mathematical Morphology*, *median* and *low pass filtering*. Dual *processing* is adopted to reduce, as much as possible, the ill-effects of any intensive and uneven illumination. The latter may introduce two problems, the first is in creating holes in the face region, which disturbs the algorithm's efficiency, and the second is in merging certain parts of the face/head with the background, which can deform the elliptical shape of the faces.

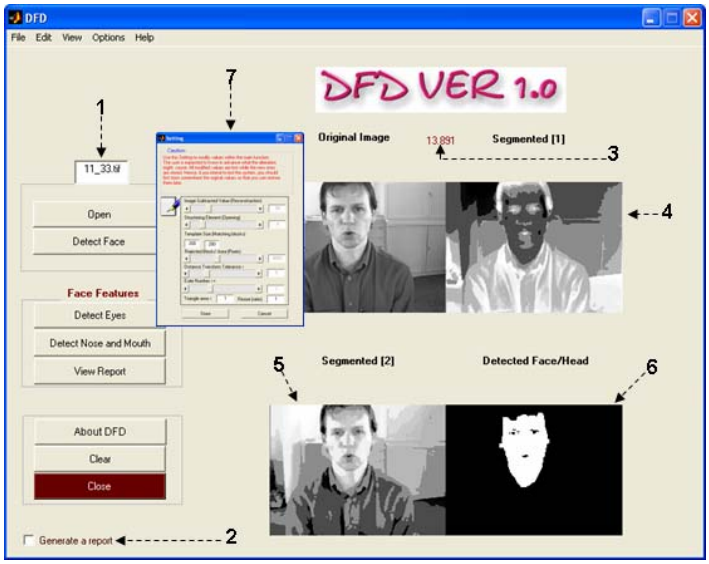


Fig. 1. The GUI of the proposed system (1) The currently loaded image, (2) An optional check box to store the process with its information in an external file (HTML) as shown in Fig. 24 in the Appendix, the generated file can be viewed or deleted by clicking the "View" drop-down-menu, (3) Elapsed time for face detection measured in seconds, (4) Image complement segmentation, (5) Direct image segmentation, (6) Segmented face displayed as a blob and (7) The settings panel.

3.2. *Image segmentation using Voronoi tessellations*

Applications of image segmentation can be found in a wide variety of areas such as remote sensing, vehicle and robot navigation, medical imaging, surveillance, target identification and tracking, scene analysis, product inspection/quality control, etc. Image segmentation remains a long-standing problem in computer vision and has been found difficult and challenging for two main reasons [27]: firstly, the fundamental complexity of modelling a vast amount of visual data

that appear in the image is a considerable challenge; the second challenge is the intrinsic ambiguity in image perception, especially when it concerns so-called unsupervised segmentation (e.g., a decision whereby a region cut is not a trivial task). A brief listing of image segmentation techniques is discussed in Section 1-1.

Our approach is based ultimately on the Voronoi Diagram (VD) technique, which employs a versatile geometric structure [28, 29, 5]. Despite its usefulness in various fields ranging from remote sensing to robot navigation, this method was largely forgotten in applications in image segmentation until [24] rekindled interest in it. The approach was based on extracting feature points which were then used to define the centres of the Voronoi cells' for image segmentation. Although our method is inline with them in using VD, our methodology and image targets are completely different. Concise definitions of VD and DT are presented below:

Given a set of 2D points, the *Voronoi region* for a point *Pi* is defined as the set of all the points that are closer to *Pi* than to any other points. More formally we can say:

Let $S = \{P_1, P_2, \ldots P_n\}$ be a finite subset of R^m and let d: $R^m \times R^m \rightarrow R$ be a metric. We define the Voronoi region VR (P_i) of a point P_i via VR (P_i) = $\{P \in R^m \mid d(P, P_i) \leq d(P, P_j)$ for all $j = 1, 2, \ldots, n, j \neq i\}$, i.e., VR ($P_i$) is the set of all points that are at least as close to P_i as to any other point of S. The set of all n VR is called the Voronoi Diagram VD(S) of S [28].

The properties of Voronoi polygons draw many researchers' attention. As a result, various algorithms have been developed to compute the VD. Among these methods we make special mention of the Divide and Conquer Algorithm, Region Growing Algorithm [29] and Fortune's Algorithm [4], and with what Reference [30] described as the *Fortune Plane Sweep* algorithm**.**

The *dual* tessellation of VD is known as Delaunay Triangulations (DT) (as shown in Fig. 2). The duality comes about in the following way: vertices in the Voronoi diagram correspond to faces in DT, while Voronoi cells correspond with vertices of DT.

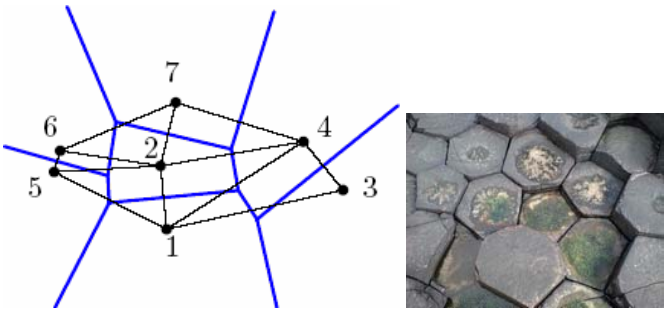


Fig. 2. (Left) Delaunay triangulation of seven dot patterns (black) and their VD (thick lines in blue). (Right) This naturally created Voronoi Diagrams on these rocks might have been the inspiration behind the discovery of VD.

Once the VD for a set of points is constructed, it is a simple matter to produce DT by connecting any two sites whose Voronoi polygons share an edge. More specifically: let **P** be a circle free set. Three points, *p, q* and *r* of **P** define a Delaunay triangle if there is no further point of **P** in the interior of the circle which is circumscribed to the triangle *p, q, r* and its centre lies on a Voronoi vertex; Fig. 3 depicts this notion.

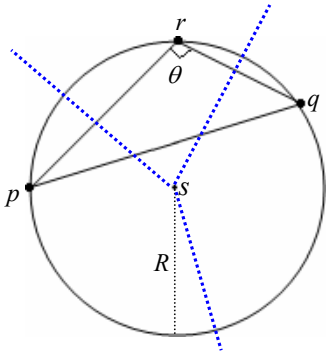


Fig. 3. Delaunay construction. *pqr* defines a Delaunay triangle when the centre (*S*) of the circle circumscribed to *pqr* is a Voronoi vertex.

Before attempting any operations, we perform a histogram equalization to reduce the lighting ill effect. This is referred to as pre-processing step. Next, the Voronoi Diagram (VD) is applied. Various literature studies have tended to apply VD on the image itself (after binarizing it and capturing its edges). This is usually time consuming; therefore, our idea was to apply VD not on the images but instead on a few selected points (<=255). Thus, VD is constructed from feature points (generators) that result from gray intensity frequencies, see Fig. 4 and Fig. 5.

The proposed technique resembles the dynamic thresholding method used for segmentation, but it differs in terms of divide and merge-decision making. Human faces have a special colour distribution that varies significantly from the background. VD is used to construct Delaunay Triangulations (DT) from points on the image histogram. The outer boundary of DT is simply the Convex Hull (CV) of the set of the feature points, as shown in Fig. 6. Hence, the two global maxima are obtained by extracting the top two values in the DT list of vertices; which correspond to the peaks in the histogram. In order to get the minima that fall between these two peaks, a new set of points are spawned using the following steps:

- Generate the host image histogram
- Generate VD /DT to obtain the list of vertices and get the two peaks
- Set all points below the first peak to zero, and set all points beyond the second peak to zero
- Set all points that are equal to zeros to be equal to the *argmax (peak1, peak2)*
- Derive the new points using:

$$Val_{new}(x) = \left| \left(Val(x) - \max\left(Val(x)\right)\right) \right| \quad (2)$$

Where *max (Val(x))* denotes the highest frequency in the host image histogram. This process will yield a local flip effect on the image histogram.

The previous course of action will expose the minima points to be part of the convex hull constructed by DT, as shown earlier in Fig. 4. These unique feature points are then sorted in ascending order to outline a one dimensional vector (V) containing values that form the ranges with which merging and splitting decisions are made. The following paradigm is applied:

```
Input: Host gray scale image Ψ
Input: Vector generated by the previously described procedure V
Initialize a new vector d = []
// The vector 'd' stores a few gray values that represent our
// segmented image. It is used to call homogenous areas that
// share the same gray value to be a binary input for our face
// localization step (best elliptical shape).
Set all pixels in Ψ smaller than V(1,1) to black (0)
Set all pixels in Ψ greater than V(length(V(:)),1) to black (0)
for i = 1 to length(V(:))−1 do
     if i ==1 then
     set all ( Ψ >= V(i) AND Ψ <= V(i+1) )
         to  V(i+1)
         d= [d ; V(i+1) ]
     else
     set all ( Ψ > V(i) AND Ψ <= V(i+1) )
         to  V(i+1)
         d= [d ; V(i+1) ]
      end if
end for
Output: Segmented gray scale image Ψ_Segmented
```

Bear in mind that the outer layer of DT contains vertices which are but some of the selected Voronoi (VD) generators, meaning that the vertices forming the *CH* are what the study is interested in. In this case these vertices give a direct access to their respective gray values from the studied image, as shown in Fig. 6.

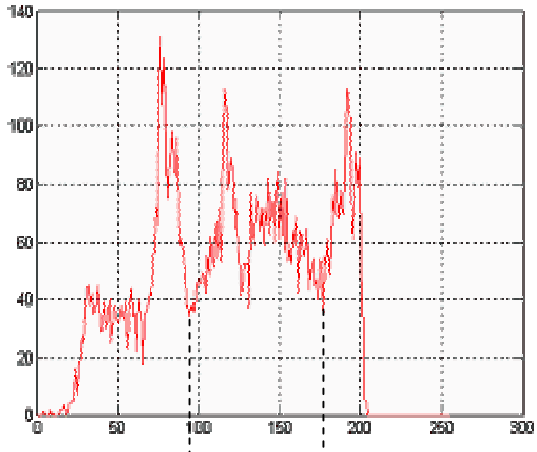


Fig. 4. Gray-level frequency (getting global maxima).

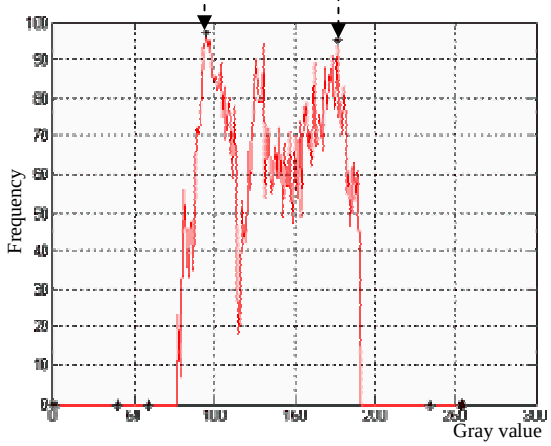


Fig. 5. Exposing local minima.

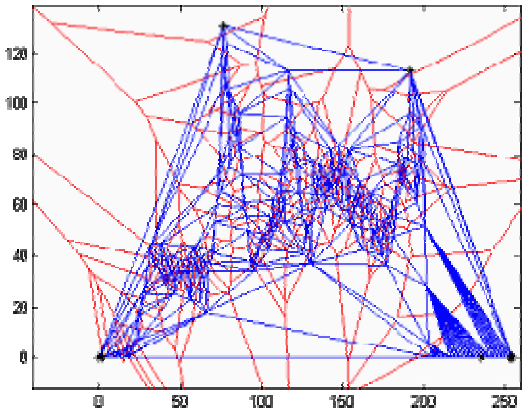


Fig. 6. Delaunay Triangulation generated from points of the histogram shown in Fig. 4.

To conclude this sub-section, the author stresses that although the scope of this research is bounded by face localization in gray scale images, the designed system can handle segmentation of other types of images, as is clear from the next example shown in Fig. 7. A second fact is that our Voronoi-based image segmentation can be extended to RGB images without converting them into grayscale. That is done by working on the RGB map only and treating each colour vector as a separate input to the segmentation process.

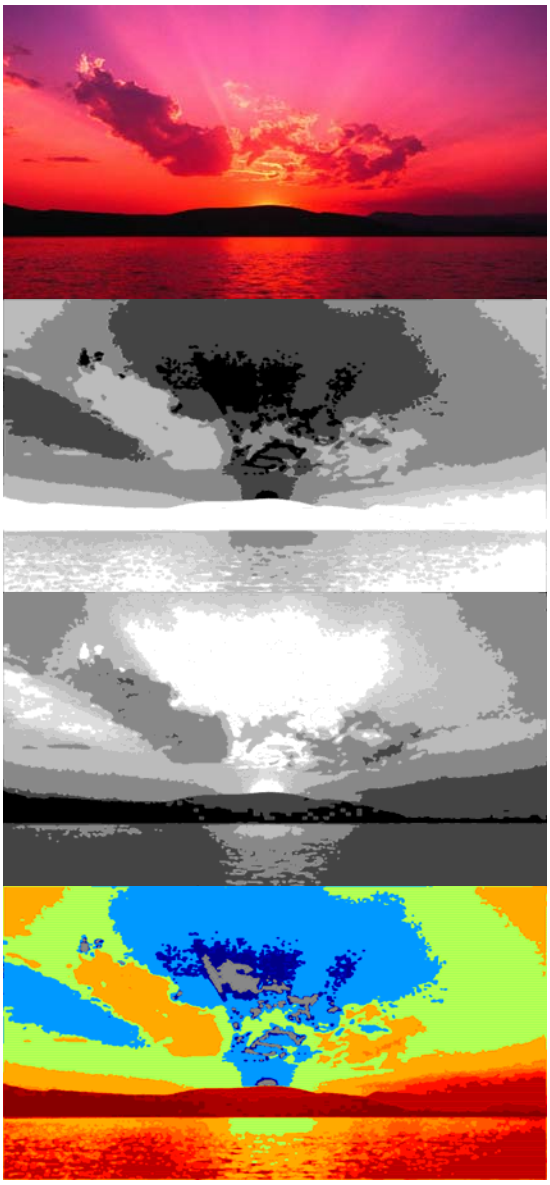


Fig. 7. Application of the algorithm on a scene of nature, as it is clear our method can be extended to segmentation of non face images. Top to bottom: Original image, direct image segmentation (DS), image complement segmentation (CS) and the corresponding colour mapping of DS.

One of the most efficient segmentation techniques is the so called K-Means clustering. For the purpose of evaluating our proposed method of segmentation, a comparison is carried out. Figure 8 shows the output of each algorithm. Unlike the K-Means algorithm, our method for segmentation does not need any parameters to be set; rather it is a fully automated and unsupervised process. All it needs in order to function is an image histogram, which is obviously an advantage.

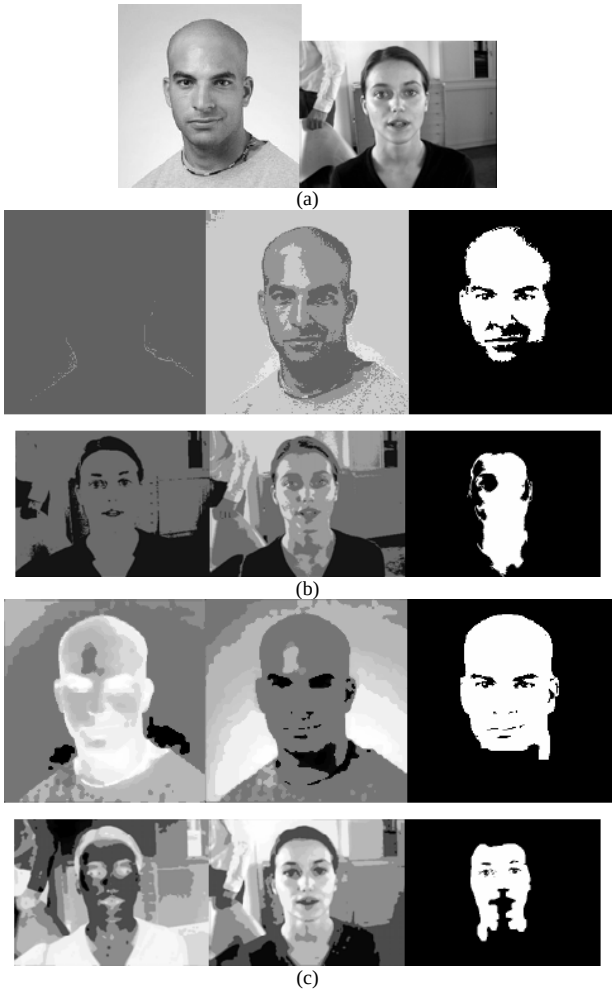


Fig. 8. Comparison of image segmentation using the K-Means and of our method. (a) Original images, (b) (Left column) output of K-Means segmentation with Gauss law, (Middle column) output of K-Means segmentation with Rayleigh law and (Right column) face region, (c) (Left column) The proposed image complement segmentation, (Middle column) direct segmentation and (Right column) face region.

3.3. *Extraction of probable face region*

A brute force is applied to each range to yield labelled connected regions. When filling all regions' holes in the image, a total (or majority) of the skin region (face and neck) will be exposed. Small regions are of no interest to us.

It has been found that faces present elliptical shapes with minor variations in eccentricity from one person to another. Thus, the voting phase can be made to be based on the detection of ellipses. The major problem researchers encounter occurs when the contrast between intensity values in face region and the values in the background surrounding the face is low. This obstacle has been reported clearly in [16, 2] while dealing with skin tone detection in colour images. According to [16], if an input image contains a background whose colour is similar to the colour of the true skin region, the labelled segments in the image can be of any arbitrary shape and need not necessarily be elliptical; therefore the hypothesized face location is not reliable at this stage. This fact was also admitted by [31] when they said: "…when the intensity difference between background and skin-region of the face is small, it is difficult to correctly extract the face region …". To cope with such situations, in this case we have used the Distance Transformation (*DisT*) to separate the face region from its background by setting up a suitable threshold value for the *DisT*. The function operates on the Euclidean distance between two given points (`x1, y1`) and (`x2, y2`), where:

$$DisT = \sqrt{(x1-x2)^2+(y1-y2)^2} \quad (3)$$

It works fine in binary images, given that the bridge connecting a face with its background is not large. A large bridge is a very rare case because the presence of head hair prevents that from happening and the merging – if it occurs – will be in the region of face sides, as depicted in Fig. 9. Three main criteria are used for estimating and evaluating the Region Of Interest (ROI), which are:

- Ellipse fitting: This is a well-known strategy for face estimation and segmentation. The connected components given in the above stage will be fit to an elliptical model. An ellipse is ultimately characterized by its centre (X_o, Y_o), orientation (θ), minor axis (a) and major axis (b). The centre is the centre of mass, while θ is determined using the least moment of inertia [3].
- Aspect ratio of the major and minor axes: This somehow is a complement to the above, where we select face candidates based on:

$$\forall a / b \leq \xi \quad (4)$$

  where ξ is a tolerance factor and *a* and *b* are described earlier.

- Euler number: A binarized face region will normally expose face features as holes because of their low intensity. This helps in identifying a probable face region in terms of its Euler number. Therefore, it is safe to say that a face would consist of at least two holes.

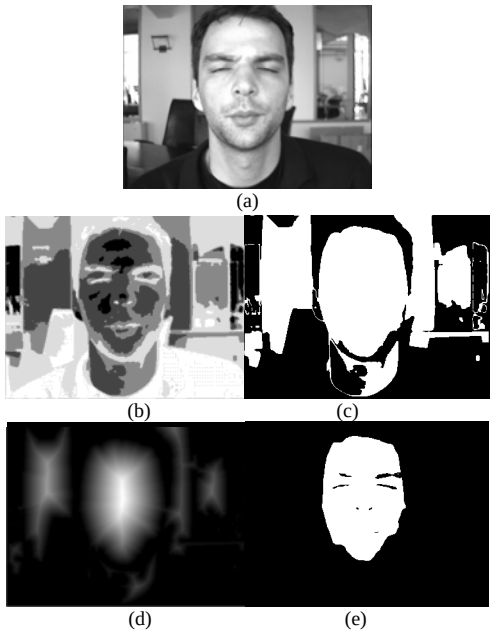


Fig. 9. Applying Distance Transformation to separate the face when it is merged with the background. (a) Original image, (b) Voronoi-based image segmentation, (c) face boundary merged with the background, (d) the Distance Transformation of (c) with an appropriate threshold, and (e) the final segmented face.

A dual processing scheme is applied to get the most likely segmented face region. The algorithm discussed in Section (3.2) will be applied to the image and to its complement in order to compensate for the drawback of lighting ill effects. The proposed method is another way to represent an image with reduced intensity values, as is clear from the experiment made on one of the BioID images (Table 1). The re-sampled result has the minimum gray values while preserving objects shapes in the image. This helps in reducing the time to search for the best fit ellipse in these objects. A final example of segmentation is depicted in Fig. 10.

Table 1. The Benefit of Segmentation in Reducing Colour Representation (Re-sampling).

Gray Values in the Host Image	*Gray Values after Segmentation using our Method*	
	Direct	Complement
252	20	19

Fig. 10. Test results of face segmentation on Set 3 (BioID). (Columns left to right) Original images, Voronoi based image segmentation and segmented faces.

As can be appreciated in Fig.16 and Fig. 17 in the Appendix, certain factors, including uneven face illumination and occlusion, can demonstrate holes when the face is segmented. A problem occurs when this happens that affects the most valuable feature for our system and for the human beings (i.e., eyes). To cope with this situation, a face boundary repairing method is introduced. The Convex Hull is called again to surpass this difficulty and to bring back the "lost portion." The following figure (Fig. 11) shows the face blob before and after repairing.

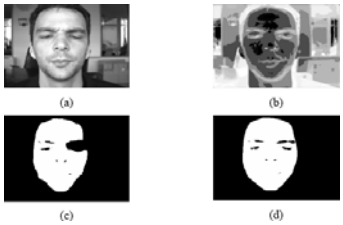


Fig. 11. Face repairing. (a) Original BioID image, (b) image segmented into homogenous regions, (c) face blob detected but the left eye region is missing, and (d) face boundary repaired thanks to the convex hull.

The ellipse model that we developed (shown in Fig.14) can adapt to such different situations as changes in rotation and scale. The only critical issue in this process is to detect eyes accurately; the ellipse will then use the measurement of the distance between the two eyes as a parameter for its own construction, taking into account the object rotation angle and scale factor.

4. Face Feature Extraction

This section is devoted to discuss *face feature extraction*. Although different researchers have different views on which features to consider, the author emphasizes extracting of the eyes, nose and mouth because of their saliency. Extracted features play a very important role in the recognition phase. Therefore, one must be very careful in choosing such features. Usually misclassifications are due to how well these features are extracted.

4.1. *Detecting eyes*

Eyes come at the top priority in the hierarchy of features. They demonstrate the following characteristics, which make them attractive:

- Eyes are identical pairs with a special distance from face sides.
- Extraction of eyes automatically leads to determination of face orientation and approximates its size.
- Extraction of eyes automatically eases the process of extracting other features (e.g., mouth, nose, eyebrow, etc.).
- Eyes by themselves are features, and the distance between them is another very important feature that is extensively used in criminal investigation at police stations.
- Eyes are characterized by the phenomena of *"Eyeness"* which is a "W" shape in the projection space.
- The composition of iris, pupil, eyeball, and eye lids give a unique description to eyes.

Given the sequence of coordinates of Delaunay triangles (e.g., Possible feature points), the pseudo-code described beneath follows:

```
Input: Original image Ψ , possible feature points, eye binary template model (TM)

Construct: Voronoi Diagram and derive the Delaunay Triangles

For i =1 :length(Triangles_matrix) do
    Calculate the area of each triangle using Hero's formula

        If area > 4 then
            Delete (Triangle)
          //Deletion of noisy points or the ones outside the cluster
        End if
End for

Use the remaining triangles to form clusters*
For each cluster do
    Combine it with each of the other clusters simultaneously
    Calculate the corresponding correlation with (TM)
    Get the maximum likelihood of a pair of eyes
End for

For each of the two eyes blobs do
     Fill the blob's pixels with their gray values from Ψ
     Search for the darkest pixel's coordinates
End for

Output: The two eyes coordinates (x , y)
```

***** A cluster here refers to dense dot patterns that belong to a single object, as can be seen in Fig. 12.

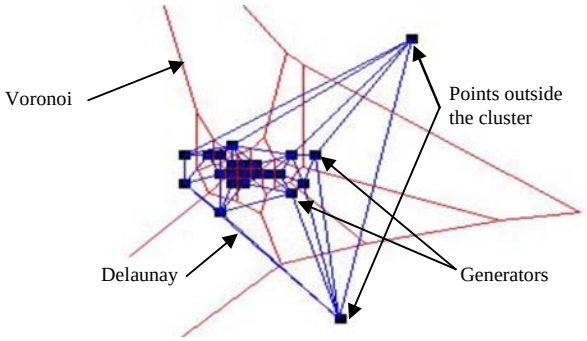


Fig. 12. Clustering concept based on Delaunay Triangulations.

The coordinates of both eyes are obtained by searching for the darkest pixel in each eye blob (i.e., pupil marks the centre of the eye). This algorithm is invariant in respect to scale, rotation and translation. The lighting condition in this stage does not affect the process's efficiency because eyes exhibit the darkest regions. For some example outputs of the algorithm on different set of images, please refer to Fig. 17 - Fig. 22 in the Appendix. Table 2 shows the success rate for

eye detection. The test results shown in tables 2, 3 and 4 were conducted on computer-generated random images from the BioID database, as discussed earlier.

Table 2. Detection Rate of the Eyes.

	Average Time in Seconds per Image	*Successful Detection Rate of Eyes*
The Proposed Method	1.9070 s	90.2%

4.2. *Detecting nose and mouth*

The mouth has almost the same importance as the eyes in face recognition. Given the positions of the eyes, one can determine that the mouth can lie only within specific areas in the image, within a predefined distance from the identified eyes. In this case, the geometrical face model is applied to locate facial features. The model utilizes the distance measurement between the detected eyes. The author estimates that in most of faces the distance between eyes and mouth is proportional to the distance between the two centres of the eyes. Let the distance between the two centres of the eyes be D, then the geometrical face model and its relative distances can be described as follows.

1. The vertical distance between two-eyes and the centre of mouth is D.
2. The vertical distance between two-eyes and the centre of the nostrils is $0.6D$.
3. The width of the mouth is D.

A variety of published researchers have responded favourably to this method [32, 33]. However, there is still a need for computation to approximate the location of the mouth. Therefore; an ellipse constructed from the known measurements is proposed. Figure 14 (Appendix) shows a model for the proposed ellipse. An ellipse accepts four parameters for its construction; herein in relation to eye distance, are defined as:

- Centre of the ellipse (X_0, Y_0): Is the centre of the distance between the two eyes?
- Minor axis length *(a)* : Is the distance between the two eyes where both eye centres lie on each side of the ellipse
- Major axis length *(b)*: Is $2D$ where D denotes the distance between the two eyes
- Angle $(\angle)$: The ellipse must have the same orientation as the detected face. We can easily determine a face orientation based on the angle made by the baseline (a line connecting both eyes) and the horizontal x axis.

The next step is to calculate the Euclidean distances from the centre of the ellipse to its border points (in a radar scan fashion), searching for the two maxima. These two maxima can approximate the position of the mouth and give us two areas where we can correlate them with a mouth template to find the best likelihood.

Certain cases have been examined where the mouth, or part of it, were out of the region of the segmentation blob; in this case the cross correlation will make no sense. For this special case, assistance comes from locating another feature, namely the nose, which, when paired with the eyes, suggest one location from the two probabilities that the mouth generated earlier.

Another geometry measurement is followed here to locate the probable nose area. It is the same procedure as the one for the mouth described earlier in this section but it differs in terms of the length of the major and minor axes:

$$(a=b= (D/2)) \tag{5}$$

Where D denotes the distance between the two eyes.

This means that the algorithm generates a circle instead of an ellipse. It is the circumscribed circle which is the largest circle bounded by the ellipse. Seven unique points are generated (i.e., two eyes, centre of distance between the eyes, two possible nose areas and two possible mouth areas), as shown in Figure 13. These unique points will help in isolating the features from one another and give each its respective size by using distance transformation and watershed methods. Examples are given in Fig. 22 in the Appendix.

Since there are two possible points with the same geometry towards the eyes for both mouth and nose, there will be no choice but to use the information from the original image to correlate for the best nose match. Hence, when one location of the two is voted on to be the nose, the elimination of one of the points generated for the mouth location will be straightforward (i.e., selecting the nearest mouth blob to the detected nose from the two probabilities) and hence the mouth location will be found. The detection success rate for nose and mouth features was high,

thanks to the previous steps. In Table 3, the percentage of correct detection is shown. It can be noted that the detection rate for both the nose and the mouth is the same; this indicates that these features are interrelated.

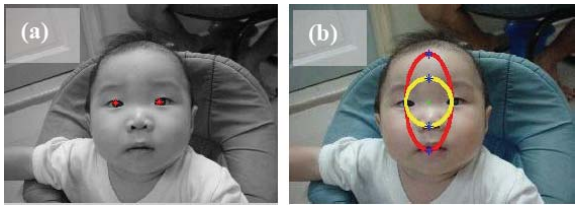


Fig. 13. An image from WWW, our ellipse and circle model superimposed on the detected face. (a) Eyes detected and (b) Geometrical distance measurement generating seven unique points.

Table 3. Detection Rate of the Nose and Mouth Blobs.

	Average Time in Seconds per Image	*Detection Rate of Nose and Mouth*
The Proposed Method	3.6250 s	98.03%

The coordinates of the eyes, the mouth and the nose are very useful parameters in the recognition process (see Table 5). They can feature a face by themselves or by feeding them into other descriptors (Fourier transforms, Voronoi Diagram, Fuzzy logic process, etc). However, sometimes a need arises to preserve the intensity values of the region of these features rather than just the coordinates, as shown in Fig. 19, Fig. 21 and Fig. 22 in the Appendix (e.g., to be used for Normalized Cross correlation, Gabor wavelet, Moments, etc).

4. Comparison to other Systems

We carry two types of comparisons of current systems encompassing either face localization or recognition. We do not claim that our work is the earliest to use VD/DT in biometrics; however, our unique contribution lies on their application on image histogram (1D data) rather than on the 2D image data which was the choice adopted by all VD/DT based methods we cite in the next section.

4.1. Other VD/DT based Face Detection/ Recognition Systems

Xiao and Yan [5, 35] used a *Constrained Delaunay Triangulation* (CDT) on the exterior contours of the blobs in a binarized facial image. They showed that three types of Delaunay triangles will be generated in such scenario. The facial features are identified by the symmetry related distance analysis and then the face boundaries are traced with a group of connected Delaunay edges. They illustrated their experiments on a bunch of images (304) selected from the set of ORL database which has faces occupying most of the image's space, as such their algorithm seemingly show some feasibility since it encountered less complex and noisy background. Fig. 25 reveals the notion of CDT. Their correct location rate reached to 89%.

A model based method was recently introduced by Xie and Lam [36] who benefited from the manifold structure that faces demonstrate. They employ an *Elastic Local Reconstruction* (ELR) technique to measure the similarities between the image block pairs in order to determine the difference between the two face images. They overlaid the face candidate by an updated version of the CANDIDE[2] model which is a parameterized face mask specifically developed for model-based coding of human faces. CANDIDE uses a number of polygons which can be viewed as triangles or vertices. Although there was no assertion pointing to the use of VD/DT but the topology of CANDIDE resembles to DT that can achieve the same results. Their work exploited CANDIDE for recognition purposes under various facial expressions of which they show an extensive experiment on different databases.

Kakumanu and Bourbakis [37] demonstrated an interesting work in which they proposed a novel facial expression recognition method based on Local-Global Graphs (LGG). The LG Graph embeds both the local information (the shape of facial feature is stored within the local graph at each node) and the global information (the topology of the face). Facial expression recognition

[2] http://www.icg.isy.liu.se/candide/, accessed on 14-04-2008 at 14:50

from the detected face images is obtained by comparing the LG Expression Graphs with the existing LG expression models present in the LGG database. Each facial feature region forms a node which is connected to other neighboring features by means of Delaunay Triangulations (DT). This topology is used for their aim to provide an invariant recognition system. The method is explicitly described in [38]. Since in their work face candidates are extracted by neural based skin tone detection therefore the algorithm works on colour images. After the extraction of the face region, it has to undergo a careful colour segmentation, which was another interesting contribution, to emphasize the facial features. Fig. 26 shows the Delaunay Graph. Unfortunately, their system assumes that the size of the face is more than 96x128 pixels and the face is always frontal.

In [39], Ersi and Zelek presented a novel technique for face recognition which represents face images by a set of triangular labelled graphs generated by salient points derived by applying Harris corner detection method. Each of the graphs contains information on the appearance and geometry of a 3-tuple of face feature points. Their method automatically learns a model set and builds a graph space for each individual. They tested the algorithm on FERET database where 97.7% correct recognition rate was achieved.

Faces can be warped using an elastic representation of face model to cope with the different changes facial expressions exhibit. DT was the way to realize that as discussed in [40, 41, 42]. In [41]; however, the topology was extended to handle 3D face models.

4.2. Other Face Detection/ Recognition Systems

Tsalakanidou et al., [43] argue that incorporating 3D range data (depth images) can boost the performance of 2D face recognition systems. They utilised a low cost 3D sensor to produce the aforementioned depth image.

Song et al., [44] take a bottom-up approach. They found that their *Binary Edge Image* (BEI) is better generated in the wavelet decomposition since edges do not suffer from having broken segments, hence preserving shapes' concavity. The ease of detection of reflected light spots on the eyes was exploited to locate faces. They ran their experiments on Bern database where they obtained 98.7% correct eye detection rate and 96.6% on AR face database. A drawback of their system, which was highlighted in their work, occurs when there is too much hair on the two sides of the face that distorts its boundary.

Xu et al., [45] introduce AdaBoost algorithm to locate roughly the pair of eyes. It was followed by a fine detection using *Minimum Extreme Region* (MER). They obtained 61.9% correct eyes detection rate. This low rate is due in part to the complexity of the test data (i.e., BioID database) which was avoided by most of the works we previously cited.

Ramirez and Fuentes [46] propose combinations of different classifiers including: *Naïve Bayes* (NB), *Support Vector Machine* (SVM), *C4.5 rule induction* and *Feedforward Artificial Neural Network* (ANN) to vote for the best face candidate. This complex system ended with high detection rate on BioID database of 95.13% but coupled with 7075 false positive face instances.

Hamouz et al., [47] start with the following assumptions, faces are flat objects apart from the nose, in-depth rotation is small and faces are close to frontal. They use a bottom-up scenario where they aim to detect 10 points on the face using Gabor filters and appearance model in the final stage. They obtained a modest rate of correct eyes detection of 73.4% on BioID database. They admit that BioID database is regarded as a difficult data set, which reflects more realistic conditions. Each image took them 13 seconds, in average, to process.

4.3. Our System

Herein we provide a summary of our method in light of what have been mentioned earlier (see, section 4.1). Our VD/DT implementation is twofold, first we exploit it in the initial image segmentation phase via image histogram (≤ 255 points to process), obviously this step has reduced significantly our search for the face best candidate, and the second use of VD/DT is on clustering of face feature blobs which shows a better constellation of local facial features. Moreover, our proposal is invariant to rotation (see, Fig. 27). Unlike [37] our algorithm is able to handle small instances of faces as can be clearly judged by looking at Fig. 28 and Fig. 29. The distance transformation applied on the obtained face blob, with seven points representing possible face features, led the process to efficiently segment the facial features regardless of the face scale. This phenomenon is shown in Fig. 28 and Fig. 29. Finally, a comparison of the

performance of our algorithm versus other published work is highlighted next in Table 4; note that we used the BioID face database to generate all the tables cited in this work.

Table 4. Comparison with other systems, (**) Denotes the number of processed windows.

Face Detection Approach	*Average Processing Speed Measured in Seconds for each Image*	*Success Detection Rate (%)*
[34]	20.5 s	94%
[33]	170 s (2min 50s)	Not reported
[8]	37.9s + 26.7s = 64.6s	87.1%
[18]	408953 PW** = 2100s	90.7%
The Proposed Method	16.7 s	95.14%

5. Conclusion

We proposed an algorithm for extracting a face and its features from a given image. We made use of the Voronoi Diagram for face localization by segmenting the image regions into connected homogeneous regions based on their gray values. The proposed method is robust, precise and independent of translation, rotation, and scaling. However, the presence of background intensities similar to the face intensities in the image may cause problem for segmentation decisions. This was also a major problem in colour-based algorithms introduced in [14, 2]. Facial hair is not an obstacle here, but wearing glasses can split the ROI (i.e., face blob) into two blocks, which means each block will be treated as different face candidate. We believe that our algorithm can be injected into many different scientific fields; for example in robotics to recognize specific people or shapes, in criminal investigations at police stations, in surveillance tasks, in driver protection (tracking driver's eyes), in Steganography [48] and many other examples.

The aim of this work was to provide a new feasible prototype, and we note here that there was little concern given to the speed factor. Hence, the reported speed of less than 16.7s could be brought down to a fraction of a second if the algorithm is implemented to another programming language (e.g., C++, C#, Java, etc). MATLAB, despite its high reputation as a scientific tool, is admittedly slow in handling nested for loops as it is an interpreted language.

Acknowledgments

This research was supported by an IRPA (Intensive Research in Priority Areas) grant from the Malaysian Government (Vote no: 74020). The first author acknowledges the Fellowship Award provided by Universiti Teknologi Malaysia (UTM) at Johor. We extend our thanks to the anonymous reviewers for their valuable constructive criticism.

Appendix

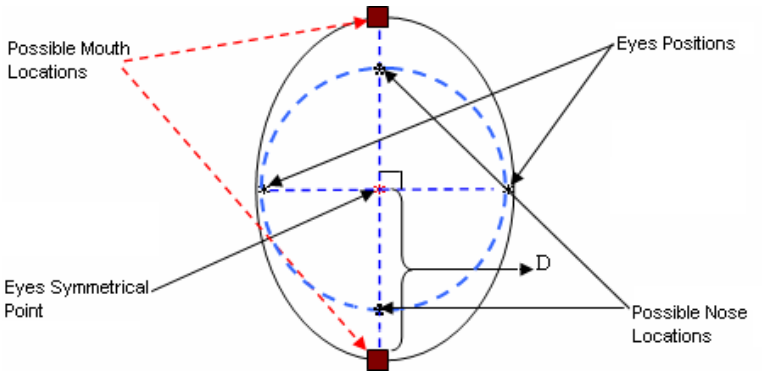


Fig. 14. The proposed ellipse and circle for mouth and nose detection.

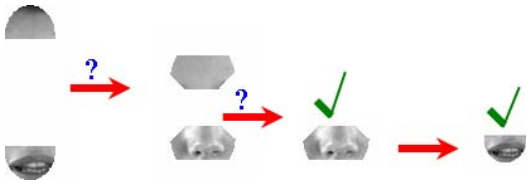


Fig. 15. The Process of mouth detection. (a) The two possible mouth regions, (b) The two possible nose regions, (c) Detected nose and (d) Detected mouth, the nearest location to the nose of the two possibilities.

Fig. 16. An example of the test set from the BioID database.

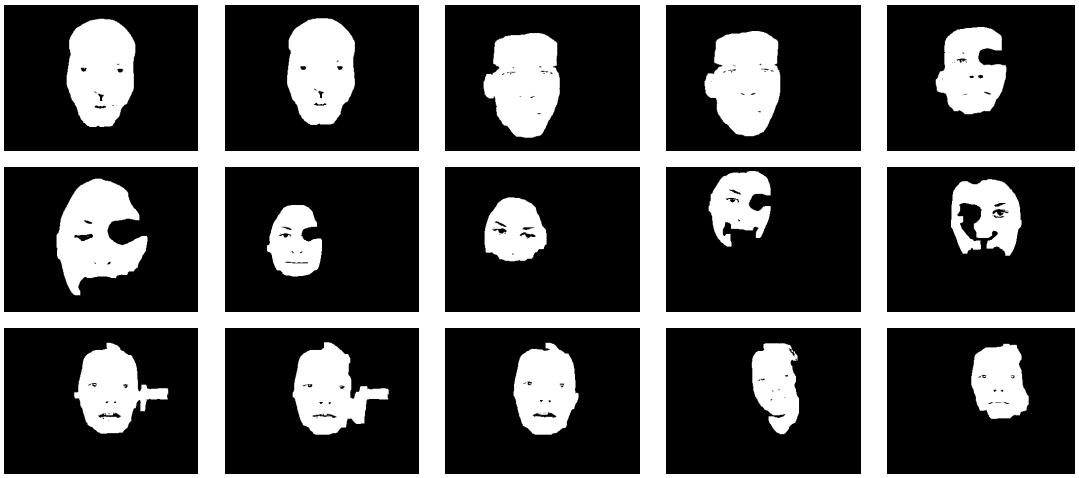


Fig. 17. Segmented faces corresponding to the set shown in Fig. 16.

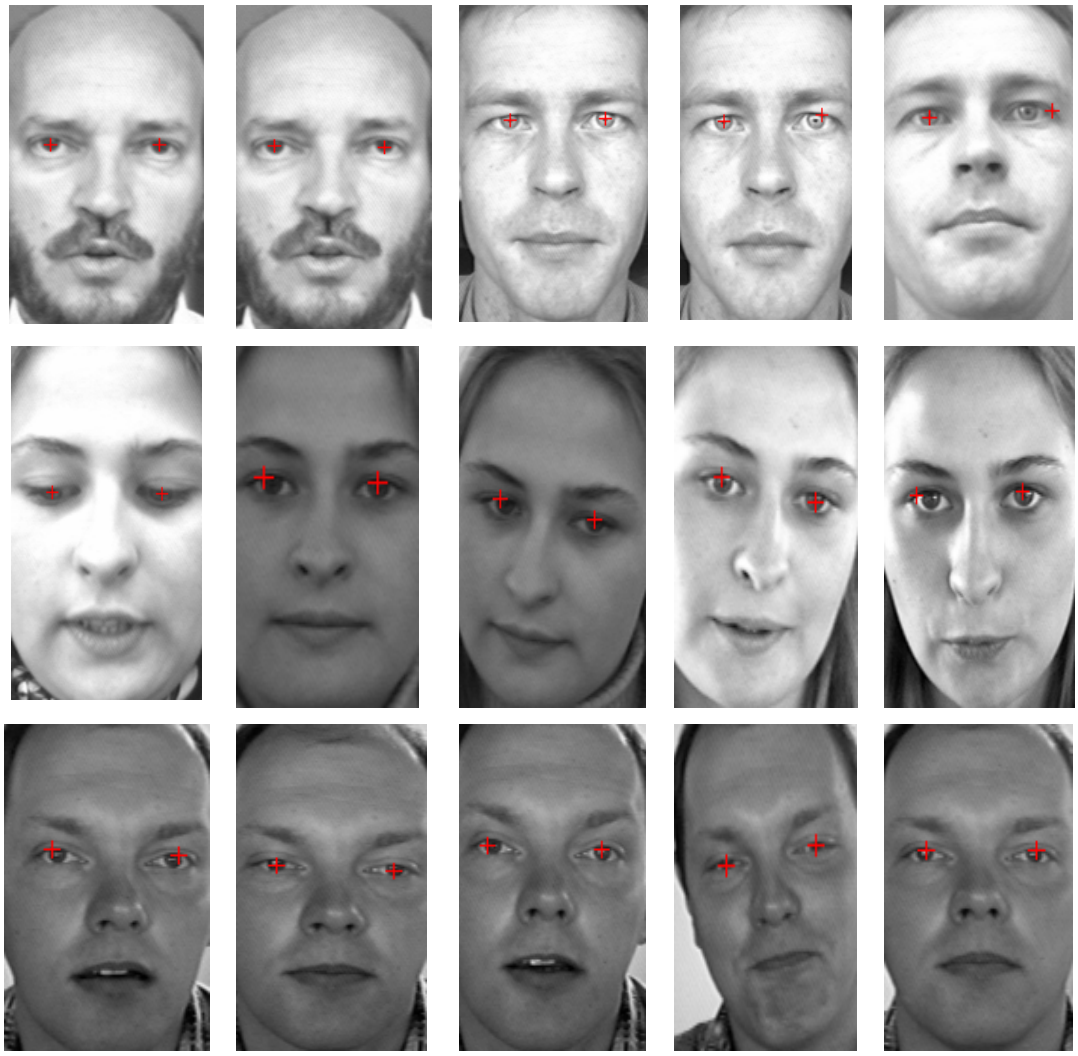


Fig. 18. Automatic location of eyes shown as red cross signs. Faces are cropped and zoomed manually to enhance the visualization.

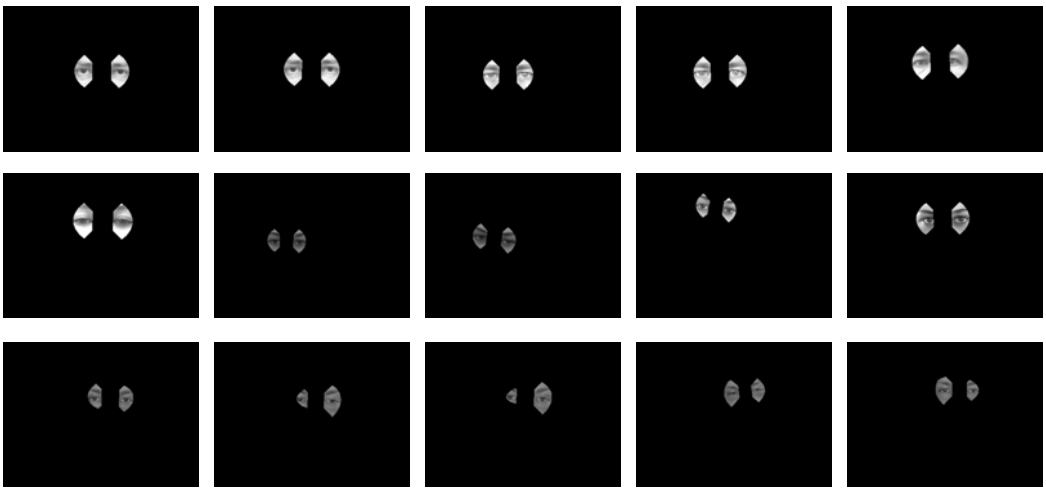


Fig. 19. Eye regions extracted. Note that the applied *DisT* gives each face feature its respective size.

Table 5. The Coordinates (x, y) of Features of the Set shown in Fig. 16 (top down and left to right).

	Image1		*Image2*		*Image3*		*Image4*		*Image5*		*Image6*		*Image7*	
Eye1	128	158	124	155	135	130	131	130	112	146	138	156	133	117
Eye2	128	230	125	228	134	196	127	200	108	220	139	237	135	168
Nose	200	194	198	192	201	163	199	165	184	183	220	197	185	143
Mouth	164	194	162	192	168	163	164	165	147	183	180	197	160	143
	Image8		Image9		Image10		Image11		Image12		Image13		Image14	
Eye1	124	108	63	131	91	153	109	177	114	169	109	164	102	191
Eye2	134	164	74	183	89	223	112	244	117	231	111	230	94	238
Nose	186	136	122	157	160	188	178	211	178	200	176	197	146	215
Mouth	157	136	96	157	125	188	145	211	147	200	143	197	122	215
	Image15													
Eye1	94	193												
Eye2	94	249												
Nose	150	221												
Mouth	122	221												

Fig. 20. Eyes detected shown as red crosses. (a) Sample from the BioID database, (b) Sample from our CCD camera and (c) Samples from the WWW.

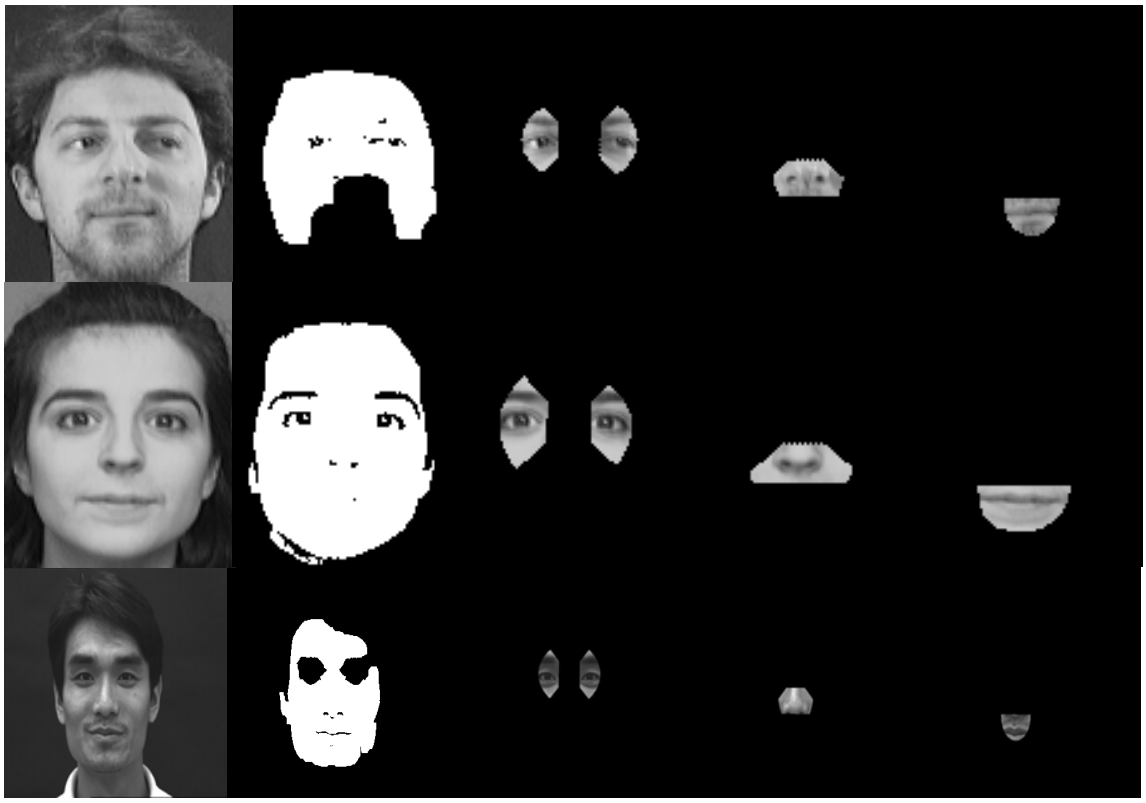
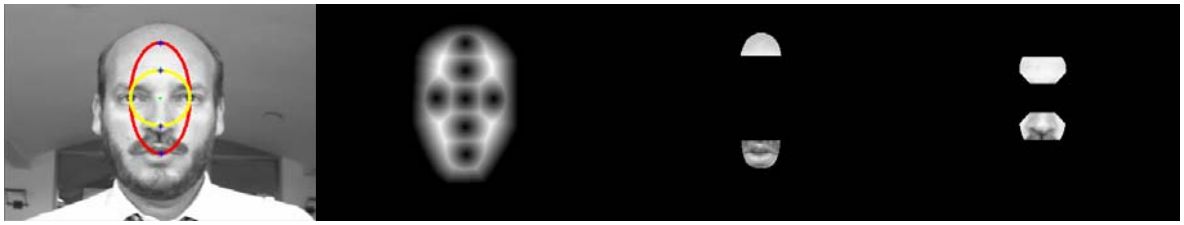


Fig. 21. Feature extraction. Top two rows refer to ORL database while the bottom one refers to our CCD database.



(a) (b) (c) (d)

Fig. 22. Mouth and nose probable locations. (a) Geometrical distance measurement generating seven unique points, (b) Distance Transform applied on the seven points, (c) Extracted probable mouth regions and (d) Extracted probable nose regions.

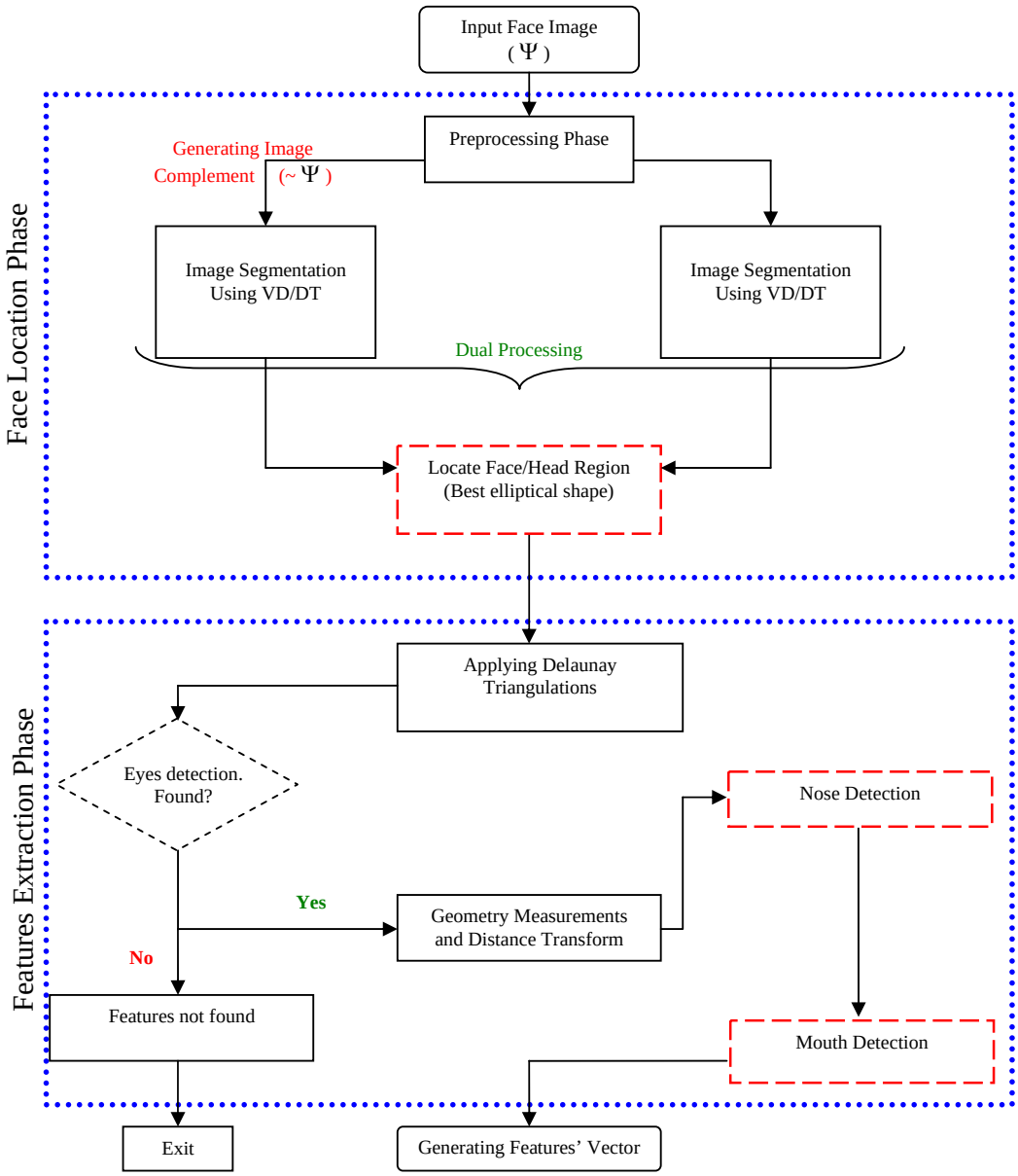


Fig. 23. A general block diagram of the proposed system.

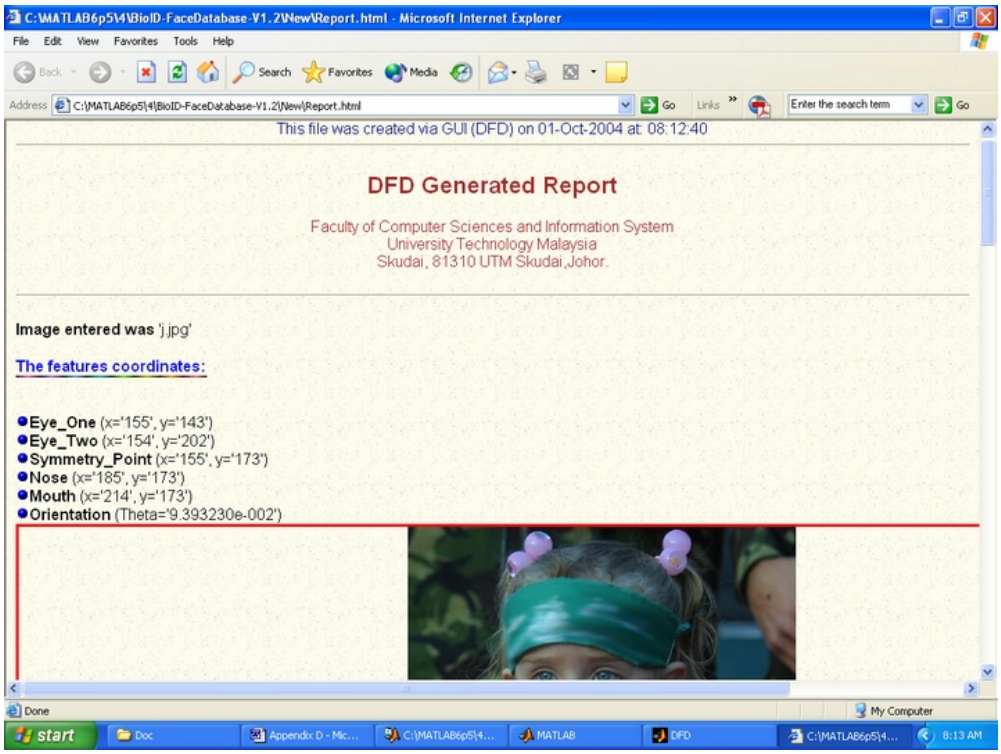


Fig. 24. The HTML report generated by triggering the check box appearing in Fig. 1.

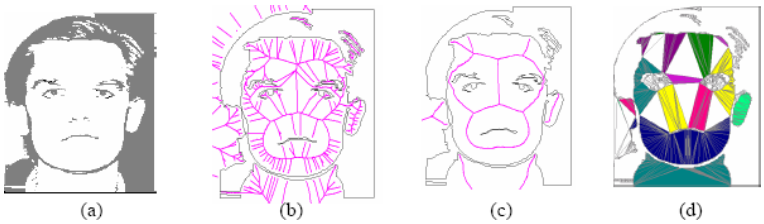


Fig. 25. Face (skin) skeleton and face shape refinement (a) A binarized face image; (b) Initial skeleton; (c) Skeleton after artifacts removing; (d) components in the face after artifacts removing. The S-triangles in each component have the same color [35].

Fig. 26. Delaunay Graph (Left) Model face image; (Middle) Skin synthesized image; (Right) Selected facial regions and Delaunay graph [37].

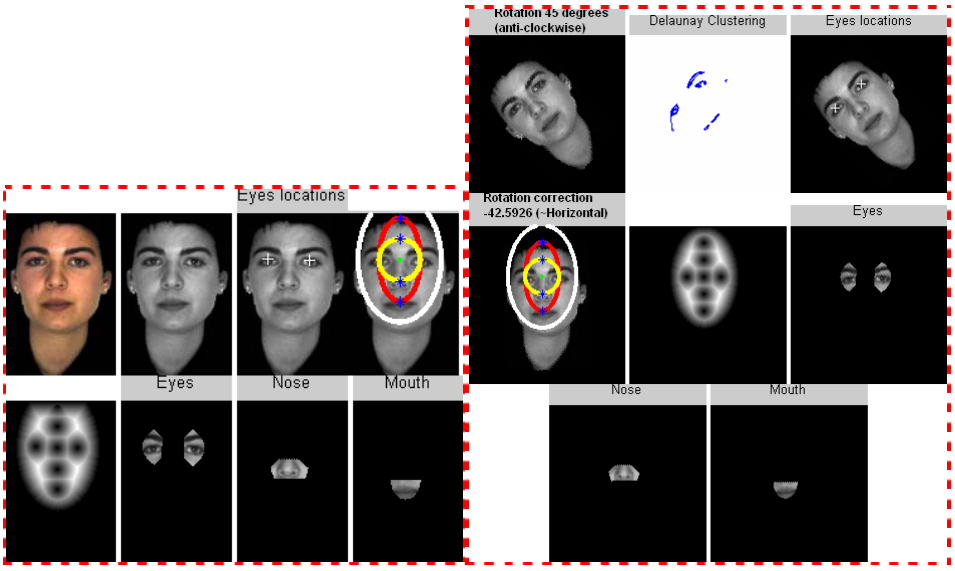


Fig. 27. Our system's rotation invariance. (Left box) features extraction before rotation and (Right box) Rotating the face model image by 45° counter-clockwise. Note, even though the rotation was approximately corrected (θ =-42.5926) but the features were extracted perfectly.

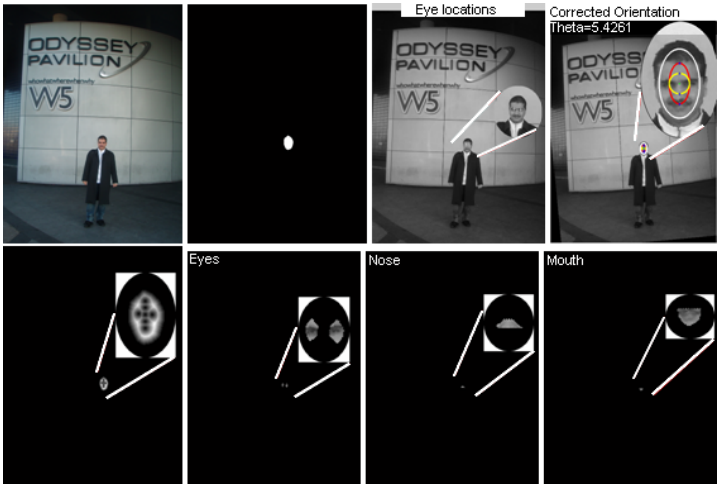


Fig. 28. Our system's adaptation to scale changes. Face size (25x37).

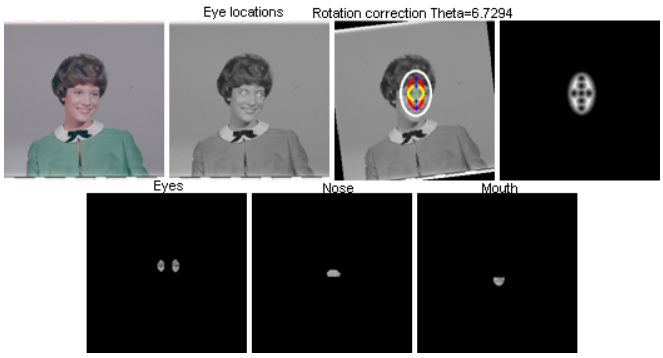


Fig. 29. Our system's adaptation to scale changes. Face size (65x88).

References

[1] K. Bowyer, (2004). Face Recognition Technology: Security Versus Privacy. *IEEE Technology and Society Magazine*: 9-20. Spring 2004.

[2] R. Hsu and M. Abdel-Mottaleb. Face Detection in Color Images. *IEEE Transaction on Pattern Analysis and Machine Intelligence*. May 2002, 24(05): 696-706.

[3] J. Haddadnia, M. Ahmadi and K. Faez, (2002). An Efficient method for Recognition of Human Faces Using Higher Orders Pseudo Zernike Moment Invariant. *IEEE Proc. Fifth Int, Conf on Automatic Face and Gesture Recognition.* 20-21: 315 – 320. May 2002.

[4] M. Seul, L. O'Gorman and M. Sammon, (2000). Practical Algorithms for Image Processing. USA: Cambridge University Press.

[5] Y. Xiao and H. Yan, (2002). Facial Feature Location with Delaunay Triangulation/Voronoi Diagram Calculation. *Australian Computer Society, Inc. The Pan-Sydney Area Workshop on Visual Information Processing*. 2002 (VIP2001).

[6] A. Somaie, (1996). Face Identification using Computer Vision. University of Bradford USA: Ph.D. Thesis.

[7] G. C. Feng and P. C. Yuen, (2001). Multi-Cues Eye Detection on Gray Intensity Image. *Elsevier Science Pattern Recognition*. 34(5): 1033-1046.

[8] V. Perlibakas, (2003). Automatical Detection of Face Features and Exact Face Contour. *Elsevier Science Pattern Recognition Letters*. 24 (2003) 2977–2985.

[9] T. Chang, T. Huang and C, Novak, (1994). Facial Feature Extraction from Color Images. *IEEE 12th IAPR Int. Conference on Computer Vision and Image Processing*, pp.39-43.

[10] H. Wang and S. F. Chang, (1997). A Highly Efficient System for Automatic Face Region Detection in MPEG Video. *IEEE Transaction on Circuits and Systems for Video Technology*. 7(4): August 1997, 615-628.

[11] S. A. Sirohey and A. Rosenfeld, (2001). Eye Detection in a Face Image using Linear and Nonlinear Filters. *Elsevier Science Pattern Recognition*. 34 (2001) 1367–1391.

[12] J. Wang and E. Sung, (1999). Frontal-view face detection and facial feature extraction using color and morphological operations. *Pattern Recognition Letters.* 20 (1999): 1053-1068.

[13] J. Dowdall, I. Pavlidis and G. Bebis, (2003). Face Detection in the Near-IR Spectrum. *Elsevier Science Image and Vision Computing.* 21 (2003): 565–578.

[14] K. Sobottka and I. Pitas, (1996). Extraction of Facial Regions and Features Using Color and Shape Information. *Proc. International IEEE Conf on Image Processing*. 1996: 483-486.

[15] Hu, S. Worrall, A. H. Sadka and A. M. Kondoz, (2004). Automatic Scalable Face Model Design for 2D Model-Based Video Coding. *Elsevier Science Signal Processing: Image Communication.* 19(2004): 421-436.

[16] B. Kwolek, (2003). Face Tracking System based on Color, Stereovision and Elliptical Shape Features. *IEEE Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance*. 21-22 July 2003 Pages: 21 – 26.

[17] J. Wang and T. Tan, (2000). A New Face Detection Method Based on Shape Information. *Elsevier Science Pattern Recognition Letters*. 21(2000): 463-471.

[18] O. Ayinde and Y. H. Yang, (2002). Region-Based Face Detection. *Elsevier Science Pattern Recognition*. 35(2002): 2095-2107.

[19] H. Rowley, S. Baluja, and T. Kanade, (1998). Neural network based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 20(1):23–38.

[20] K. Wong, K. Lam and W. Siu, (2001). An efficient algorithm for human face detection and facial feature extraction under different conditions. *Pattern Recognition.* 34 (2001) 1993-2004.

[21] Yang, M.H., Kriegman, D.J. and Ahuja, N., (2002). Detecting Faces in Images: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 24(1): 34-58, January 2002.

[22] Zhao, W., Chellappa, R., Phillips, P. J. and Rosenfeld A., (2003). Face Recognition: A Literature Survey. *ACM Computing Surveys.* 35(4):399–458, December 2003.

[23] Cheng H. D., Jiang X. H., Sun Y. and Wang J., (2001). Color image segmentation: advances and prospects. *Pattern Recognition.* 34 (2001) 2259-2281.

[24] M. A. Suhail, M. S. Obaidat, S. S. Ipson and B. Sadoun, (2002). Content-Based Image Segmentation. *IEEE, International Conference on Man & Cybernetics (SMC)*. Volume: 5, 6-9 Oct. 2002

[25] M. Burge and W. Burger, (1999). "*Ear biometrics*", in Biometrics: Personal Identification in Networked Society, A. Jain, R. Bolle, and S. Pankanti, Eds., pp.273--285, Kluwer Academic, Boston, MA.

[26] K. Sobottka and I. Pitas, (1998). A novel method for automatic face segmentation, facial feature extraction and tracking. *Elsevier Science Signal Processing: Image Communication.* Volume 12, Issue 3, June 1998, Pages 263-281.

[27] Z. Tu and S. Zhu, (2002). Image Segmentation by Data-Driven Markov Chain Monte Carlo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol.24, No.5 May 2002, pp 657-673.

[28] L. Costa and R. Cesar, (2001). S*hape Analysis and Classification*. USA: CRC Press, 2001.

[29] N. Ahuja, (1982). Dot Pattern Processing Using Voronoi Neighborhoods. *IEEE Trans on Pattern Recognition and Machine Intelligence*. 1982. PAMI-4.3. 336-343.

[30] A. Trigui, (2002). *Voronoi Diagram and Delaunay Triangulations*. Department Simulation of large Systems (SGS) Institute of Parallel and distributed Systems (IPVS) University of Stuttgart: Seminar.

[31] M. Rizon and T. Kawaguchi, (2000). Automatic Eye Detection using Intensity and Edge Information. *Proceedings of IEEE TENCON 2000*. Volume: 2, 24-27 Sept 2000 Pages: 415 - 420 vol.2.

[32] E. Saber and E. Tekalp, (1998). Frontal-view face detection and facial feature extraction using color, shape and symmetry based cost functions. *Pattern Recognition Letters.* 19 (1998) 669–680.

[33] F. Shih and C. Chuang, (2004). Automatic extraction of head and face boundaries and facial features. *Information Sciences.* 158 (2004) 117–130.

[34] C. Han, H. Liao, G. Yu and L. Chen, (2000). Fast Face Detection via Morphology- Based Pre-Processing. *Pattern Recognition.* 33 (2000) 1701-1712.

[35] Y. Xiao and H. Yan, (2003). Face Boundary Extraction. Proc. VII Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), 10-12 Dec. 2003, Sydney

[36] X. Xie and K. M. Lam, (2008). Face recognition using elastic local reconstruction based on a single face image. Pattern Recognition 41 (2008) 406 – 417.

[37] P. Kakumanu and N. Bourbakis. (2006). A Local-Global Graph Approach for Facial Expression Recognition. Proceedings of the 18th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'06).

[38] P. Kakumanu and N. Bourbakis. (2007). Detection of Faces and Recognition of Facial Expressions. NATO Security through Science Series, E: Human and Societal Dynamics Fundamentals of Verbal and Nonverbal Communication and the Biometric Issue. Volume 18, 2007, pp: 261 – 274.

[39] E. F. Ersi and J. S. Zelek, (2007). Local Graph Matching for Face Recognition. IEEE Workshop on Applications of Computer Vision, 2007. WACV '07. Feb. 2007.

[40] A. Z. Kouzani, S. Nahavandi, N. Kouzani, L. X. Kong and F. H. She, (2000).A morphing technique for facial image representation. IEEE International Conference on Systems, Man, and Cybernetics, (2) 8-11. Oct. 2000 Page(s):1378 – 1383.

[41] L. Yuan and M. L. Gavrilova, (2006). 3D Facial Model Synthesis using Voronoi Approach. International Symposium on Voronoi Diagrams in Science and Engineering, 2006. ISVD '06. 3rd July 2006 Page(s):132 – 137.

[42] L. Xiaoxing, G. Mori and Z. Hao, (2006). Expression-Invariant Face Recognition with Expression Classification. The 3rd Canadian Conference on Computer and Robot Vision, 2006. 07-09 June 2006.

[43] F. Tsalakanidou, S. Malassiotis, and M. G. Strintzis, (2005). Face localization and authentication using color and depth images. IEEE Transactions on Image Processing, Feb. 2005. 14 (2), pp: 152 – 168.

[44] J. Song, Z. Chi and J. Liu, (2006). A robust eye detection method using combined binary edge and intensity information. Pattern Recognition, 39 (2006), pp: 1110 – 1125.

[45] G. F. Xu, L. Huang, C. P. Liu and S. Q. Ding, (2007). Eye Location using Hierarchical Classifier. International Conference on Machine Learning and Cybernetics, 2007 Volume 4, 19-22 Aug. Page(s): 2193 – 2197.

[46] G. A. Ramirez and O. Fuentes, (2005). Face detection using combinations of classifiers. Proceedings of the 2nd Canadian Conference on Computer and Robot Vision, 2005. 9-11 May 2005, Page(s):610 – 615.

[47] M. Hamouz, J. Kittler, J. K. Kamarainen, P. Paalanen, H. Kalviainen, and J. Matas, (2005). Feature-Based Affine-Invariant Localization of Faces. IEEE Transactions on Pattern Analysis and Machine Intelligence. 27 (9), pp: 1490 – 1495, Sept. 2005.

[48] A. Cheddad, J. Condell, K. Curran and P. Mc Kevitt. (2008). Biometric Inspired Digital Image Steganography. 15th Annual IEEE International Conference and Workshops on the Engineering of Computer-Based Systems (ECBS'08). pp. 159-168.