# Automatic face identification

Charles Beumier, Marc Acheroy

Royal Military Academy
Department of Electrical Engineering
Avenue de la Renaissance, 30 B1040 Brussels
http://www.elec.rma.ac.be/index.elec.html

## ABSTRACT

This paper presents discussion and results in the field of automatic face identification. The implementation possibilities are presented, and the retained choices are motivated. The objective is to identify the person whose image is available from a grey-level camera. The approach is to extract characteristics that will be classified according to extracted characteristics of a database. One section is devoted to the importance of a proper acquisition method, here based on profile images. Several sections are more technical and deal with the profile extraction, the computation of the curvature and the way characteristics are derived. This is naturally followed by practical results. Finally, some perspectives are listed to let the present work be integrated in a practical application where several hundred of persons must be identified.

**Keywords:** face, person identification, profile, curvature.

## 1. OBJECTIVES

This research aims at studying the possibilities of automatic (computerised) face identification. It thus focuses on the ability of a computer program to distinguish which person is standing in the field of view of a sensor. Although our aim is to achieve full automatism, semi automatic solutions can help research by analysing the importance of unknown factors before implementation.

For the sake of simplicity, the problem of face recognition (the ability to recognise a face as a face, independently of the underlying person) is not addressed here. The sensor data is handled knowing that one and only one face is present. This can lead to obvious problems but as faces are rather singular patterns, non-face data normally imply unrealistic characteristics and can be easily discarded. Thus, although not intended by this research, face recognition is normally performed as a result of characteristic consistency checking.

All the research has been carried out keeping in mind practical considerations. The solution thus implies memory and time limitations. In the same spirit, the algorithms must be kept simple to avoid result sensitivity and maintenance difficulties. Finally, a simple camera producing 512×512 grey level images will suffice, leading to a cheap and largely available sensor. We can summarise the strategy evolution as one which has searched for easiest solutions, resulting in practical limitations (giving ideas about the whole problem difficulty) and implying radical changes in acquisition and software requirements.

The main applications of face identification include security and forensic systems. The former are used in access control, providing for a remote solution (no close contact or card) and storage of images useful for proof. The latter help the police in looking for criminals. In that respect, disguise or bad circumstances prevent the images from being appropriate for identification but a reduction of mugfile search is a precious tool to diminish influence of seen faces on witness's remembering.[1,2] By the use of appropriate features, but beyond the scope of our current research, the analysis of faces can lead to information about age, sex,[3] race, emotion[4] and visual communication (lip reading).[5,6] Pushed further, it could infer about social and professional category, health or fatigue.[7]

Several conclusions are hoped to be drawn from the different approaches considered (front view, profile and possibly 3D). First, their ability to discriminate among the persons in order to achieve identification, what is complicated by the natural similarity of faces (components and configuration). Secondly, the influence of natural modifications (temper, expression) and usual changes (hair) which reduce the specificity of facial measures by the computer and thus complicates classification. Third, the possibility to combine approaches in order to reach better results.

One important decision concerns the general method followed for face identification. This is so crucial it has received a full section of explanation. So the next section compares global and statistical (feature extraction followed by classification) approaches for identification purposes.

# 2. APPROACH

We face (indeed!) an impressively complex problem. On the one hand, the human faces look very similar. They all consist of the same parts, although small variations of them or their configurations make the faces very different for the human observer. On the other hand, a same face can vary a lot due to the subject's temper, expression and accessories (makeup, glasses, beard, etc).

On the computer side, the images of a face can undergo some more modifications due to lighting and point of view, resulting in a very large potential variation due to contrast and shading, translation, rotation and scale. As an illustration, even the human being (the best system for face identification) is disturbed when confronted to unfamiliar persons who have changed from attitude or hair. To reduce the variability of the images, we will benefit from the subject's co-operation especially for his attitude and orientation (see further) and we will impose guidelines during the acquisition of images. Anyway, variations are still a matter of concern which complicates practical issues.

There are two opposite classes of identification approaches : global matching and statistical analysis. In the first approach, the input image is globally matched with reference patterns, looking for the best valid match. Typical implementations concern correlation and artificial neural networks. The statistical analysis consists in the extraction of characteristics and their classification in order to identify the subject. Between both, we could place the model-based approach where some features are used to adapt a predefined model. See Mundy[8] for a comparison of those three recognition approaches.

In what follows, the two opposite approaches will be compared in terms of:

- the human ability for face identification
- algorithmic considerations
- available bibliography.

## 2.1. Human ability

There seems to be a complementary use of featural and global (called 'wholistic') approaches in the brain.[9,10,1] The global approach is used as a rapid indication of feeling of familiarity[9] and is obviously used at long distance where the resolution lacks. External parts (hair, outline) have a large contribution to the global impression and receive much importance for unfamiliar faces.

On the other hand, the statistical (featural) approach is used critically at short distance, analysing inner components. Familiar faces identification and communication naturally use this way as it relies on expressive parts (eyes, eyebrows and mouth). As described in,[9] the feature analysis, concentrating on each part to extract characteristics, takes much effort and longer time.

## 2.2. Algorithmic considerations

A direct matching (possibly after some transformation) of the input image with reference patterns has a very limited success for faces. It can't adapt to illumination and viewpoint changes or imposes too strict acquisition guidelines. The importance of hair or other flat regions hide localised features such as the eyes or mouth lines. As an exercise, the Kohonen map was used for face classification[11] on normalised images. The success was limited by the difficulty to realise a valid normalisation and by memory and time requirements. The conclusion of this work focused on the importance of an appropriate pre-processing.

In general, neural networks offer interesting properties such as the reusability of the code, the possibility to run it efficiently on a parallel machine and their ability to generalise. On the wrong side, it is difficult to analyse the reason of a miss and they require a large database (and thus memory) for learning, especially here because faces show subtle relevant variations and unuseful large modifications.

The statistical approach is completely opposite. Its implementation is application dependent and normally sequential, some features relying on the detection of others. This results in a fully controlled but complex program unless the knowledge sources are separated such as in the blackboard implementation.[12] By an appropriate strategy, we can hope for a quick computation. It's also possible to extract global (e. g. configuration) as well as local (components) features. Saving features only, storage space and classification time are dramatically reduced.

## 2.3. Bibliography aspects

Numerous global methods have been presented in the literature. The two famous ones are Kohonen Maps[13] and Wisard.[14]

Many papers mention the possibility to detect the position of the eyes. However, this doesn't mean the underlying networks are adapted to face identification because the eyes are very dark and localised areas, often detectable by a simple threshold.

Rather few papers concern the computer detection of features. One of the richest field of research in that respect is visual communication which aims at reducing the transmission rate of facial images ("Model-based image coding"). For this, an automatic system based on the detection of articulation points allows for synthesis at the receiving end with less than one hundred points and, from time to time, rendering information.[15,16] Related to this field is the synthesis of facial images for animation or surgery applications.[17,18]

The detection of reference points is crucial for the feature approach. It allows parts of the face to be isolated and measured. They thus give information about configuration (distance between parts) and classification of parts (nose length, mouth width, etc). For the latter, an interesting technique consists in matching deformable templates to parts of the image, in a roughly predetermined region.[19] The parameters of the model which offer the best match can be used as features for classification.

# 3. ACQUISITION

There exists a general compromise between generality and feasibility. On the one hand we would like to have a robust system, independent of lighting conditions, viewpoint and subject's attitude. On the other hand, we want to keep development and means (hardware and software) below an acceptable limit of complexity. Many decisions must be taken at the acquisition level to ensure the possibility of a practical solution. They imply acquisition guidelines.

Acquisition guidelines normally evolve during development. They are adopted to simplify the algorithms or to enable new techniques. They necessitate software modifications possibly making old acquisitions inappropriate for the new strategy. For these reasons, the databases have been currently limited (81 images for 20 persons).

## 3.1. What to acquire?

Perhaps the most important acquisition guideline concerns what to acquire. The face application can offer several alternatives concerning the underlying information to acquire (front view, profile) as well as the acquisition system (camera, colour, range finder, 2 cameras, ...).

A first database was set up with 256×256 8-bit (grey) images of front view faces. The information needed for simple and automatic identification was hard to find because hair, eyes and mouth undergo large modifications and eyebrows can be hidden by spectacles. Also nose and cheek outline detection were sensitive to illumination.

We thus naturally turn towards profile to study the discriminancy of more static facial parts such as the forehead, the nose and the chin and their simple arrangement in a 1-Dimensional curve. For this, a limited database was set up according to the guidelines described below.

In the future, colour will have to be considered to simplify detection and provide new features (hair colour, lips thickness, etc). The combination of profile and front view should also improve significantly the identification but at the price of additional hardware (or less resolution). In the same spirit, a 3-D sensor would provide for a new approach (less viewpoint sensitive, light and makeup insensitive) and would certainly simplify algorithms.

## 3.2. Acquisition guidelines

**Digitisation** Images are captured by a standard camera and digitised in a 512×512 8-bit (grey) format. Because the digitisation board hasn't square pixels, the image is resampled by a 4/3 factor to give 512×680 images. Due to important height variations among people, a larger dynamic has been devoted to the vertical direction by turning the camera by 90 degrees. This allows to keep the subject close to the camera for a maximum resolution. In the so-acquired images, the face typically takes about 200×300 pixels, a satisfying resolution. No constraint is really associated with digitisation. The programs are written to support a large range of sizes of grey images. Of course, the largest resolution the best, with a minimum of 256×256 pixels. To account for natural variation, at most two images of the same person were recorded on the same day and one week typically separated two records of the same person.

**Illumination** Depending on the camera quality, light sources must be added. In our case, two spots have been used, each one diminishing the shadow created by the other. This additional illumination increases the contrast between the subject and the background and between the different parts of the face. A careful (but easy) illumination is one key for the success of image processing algorithms.

**Background** For the sake of simplicity, the background has been chosen uniform. Many techniques can be considered to deal with other kinds of scenes. For instance, a static background can be removed thanks to a simple difference. A white background has been adopted to ease contour detection of the face. This also means only one (complete) face is expected to appear in the image.

**Subject's co-operation** The human face is certainly the most varying part of the human being. It is thus normal to expect from the subject his co-operation concerning his position, attitude and accessories. The position (distance and orientation) relative to the camera is critical. An appropriate distance ensures a minimum resolution and a repetitive orientation allows the analysis of a 3-D problem by a 2-D image. A normal attitude means a normal expression (especially as far as the mouth and the eyes are concerned) and the appropriate posture (profile or front view for instance). These constraints can seem natural to the subject thanks to particular arrangements (screen to capture the attention or display the subject's face for his own appreciation). For our purpose, people stand at a distance of 1.5 m from the camera and are asked to gaze at something. They shouldn't wear hat, scarf or glasses although this last requirement (glasses) is not critical. Mention that an automatic zoom, a 3-D acquisition system or a sequence record could help getting rid of many of these constraints but at the price of additional software.

**Extra information** It is always interesting to consider other types of information, especially if they are available with the same hardware system. In the early attempts, two pieces of black tape were stuck on the white background to give an estimate of the person's height even if the camera was moved a little. This easy and fruitful feature was soon abandoned due to its specificity. This kind of constraint should typically be avoided during development and reconsidered for the real application due to its obvious benefit.

# 4. DATABASE UTILITY

Human faces are very similar. It is certainly the social needs which explain the natural human ability to identify persons despite this similarity and the intrinsic variability (hair, expression, etc). What we hope for is small intra-class variations (different images of the same person) and large inter-class distances (different persons). The study of inter and intra classes statistics requires a large number of persons and of images per person respectively. However, the number of images are kept low during development as long as the acquisition guidelines are subject to modifications.

A first use of the database consists of the collection of features extracted by human operators. The input can be answers to questions (Are lips thick?, ...). One good reference of such a work is.[20] The expected results concern the coherence of judgements of an operator and the powerfulness of given features for identification.

A step towards automatism can be achieved by requiring the operator to specify directly on the images reference points from which to extract features automatically. Above the usual results about the usefulness of features, this would add the study of precision and independence.

The final step towards automatism would be to detect reference points (or regions) by software. Due to illumination, viewpoint and subject's attitude, this is very difficult. A field of application of such an ability is video conferencing where faces are synthesised at the receiver end thanks to transmitted articulation points in order to reduce the amount of data to be transferred.[15]

A last possible application of the database concerns the study of the human ability for face identification, an important field of research.[21] In our laboratory, the use of an occulometer allows for the study of eye movements during the analysis of an image. We hope to get information (or at least confirm opinions) about the human strategy and possibly apply it for computer identification.

In what follows, the database will help to verify the adequacy and the robustness of the developed algorithms and will offer an estimate of the classification powerfulness of information extracted by the algorithms.

# 5. PROFILE EXTRACTION

Prior to this profile study, a front view research has been carried out leading to the conclusion that most facial parts important for identification can vary to a large extent (hair, eyes, mouth). Moreover, the extraction of powerful features is far from being simple due to the many degrees of freedom. On the contrary, the profile of a face has the following advantages concerning its use for computerised face identification.

First, the profile outline, a 1-dimensional curve, possesses a rather important information content relative to the small number of points (about 200 in our case) it is made of. Algorithms will thus be simple, quick and will require little memory.

Secondly, the profile consists of many static parts, namely the forehead, nose and chin. These parts can hardly be modified even for a non co-operative subject. We thus foresee the adequacy of simple algorithms for feature extraction with no external influence or backtracking needs (due to wrong detections in a sequencial program).

Thirdly, the profile is rather insensitive to normal head rotations. Only the throat and sometimes the chin vary for important upward rotation.

Fourthly, the profile reveals many obvious reference points from which to derive features or to guide further extraction. The nose tip, chin and throat are some examples.

Thanks to the white background, faces are easily separated from the background. However, because the illumination of the spots can result in an important shading of the background (a variation of 40 grey levels between the top and the bottom of the image is typical), a simple threshold is not used. Instead, low pass filtering is applied to reduce the noise effect and a high pass filter reveals the important transitions at the face outline. Now background pixels are nearly all zeroes and growing is applied on obvious background pixels. We then get a binary image and an edge follower suffices to follow the outline. The successive points of the outline are stored to a file as $x$, $y$ pairs.

# 6. CURVATURE

The analysis of a curve can be performed in many ways. Here again, we could distinguish among global, model-based and statistical (feature) approaches. Anyway, it's a good practice to perform a transformation to get rid of variations such as translation, rotation or scale. For a curve, an interesting transformation is its curvature especially because it is shift and rotation invariant and because it allows for the detection of relevant reference points (local maxima). A template matching procedure based on curvature is used for profile analysis in.[16]

The way to compute the local curvature (the inverse of the local radius) is classically to consider the intrinsic equations of the curve, $x(t)$ and $y(t)$ versus the arc length $t$, and to compute[22]

$$K(t) = \frac{\dot{x}\ddot{y} - \ddot{x}\dot{y}}{(\dot{x}^2 + \dot{y}^2)^{3/2}}$$

where derivatives are relative to t. This was not considered here for several reasons. First, a digital curve has a lot of noise due to the digitisation grid. The derivatives (especially those of the second order) highlight that noise. Secondly, a local estimation is limited to seven possible curvature values (3 positive, 1 null and 3 negative values). Wuescher[22] invites to convolve the curvature with several gaussian of different variances. This reduces noise and makes the estimation less local but requires many computations. Thirdly, the dependence of $x(t)$ and $y(t)$ on $t$ complicates the evaluation, because $t$ is constrained to grid increments.

The proposed method considers the curvature at a sufficiently large scale. For this, three points at equivalent interval on the curve are used. A geometric estimation of the curvature was first looked for. The arc distance and the arc area were used individually with satisfying results. Finally, a geometric evaluation of the curvature was found as described below.

Consider the circle passing through the three points used for curvature evaluation (Figure 1). The objective is to compute its radius. Although this is possible by estimating the parameters of the equation of a circle, there exists a more direct (geometric) manner. From the figure, you see

$$R = \frac{a}{2\sin\alpha} = \frac{b}{2\sin\beta} = \frac{c}{2\sin\gamma}$$

because $\vec{b} \times \vec{c} = bc\sin\alpha$, we have

$$K = \frac{1}{R} = \frac{2\;\vec{b} \times \vec{c}}{a\;b\;c} = \frac{2\;\vec{c} \times \vec{a}}{a\;b\;c} = \frac{2\;\vec{a} \times \vec{b}}{a\;b\;c}$$

which is symmetric in $a, b, c$ as expected. A geometric interpretation is that the radius is the ratio of the product of the triangle sides by for times its area. Obvious verifications are the right angle ($c = 2R$) or the line ($R = \infty$). The main advantages of that solution are the simplicity of evaluation (3 distances, 1 cross product) and the availability of sign to distinguish between concave and convex shapes.

There is one problem left, the choice of the three points on the curve. The first one is the point where we want the curvature to be computed. The two other points lie on the curve on each side of the first point and at a distance depending on the desired scale. Unfortunately, the available points lie on the digitisation grid, so that the desired distance is only approximated (possibly using interpolation between two curve points). The effect is a limitation in precision of the curvature, whose amplitude can nonetheless be used as such, even for distances as small as five pixels.

## 6.1. Reference points

The first direct use of the outline is to get the top of the head involving some averaging to reduce the influence of hair. From this top, the curve is analysed horizontally to detect the throat, a very large reduction in horizontal co-ordinate. The ratio (vertical distance from top to throat) / (max horiz size of head) eliminates bad candidates. Possibly, due to bad outline or non-face image, no acceptable throat reference is found and the program exits.

From the throat, the most prominent part above it is searched for and the chin is found. The nose constitutes the next prominence, going up along the profile.

Those reference points were perfectly detected with very good precision (and after very little tuning) for all the images of the database. They also agree with their human localisation. Extracted features will largely rely on those points.

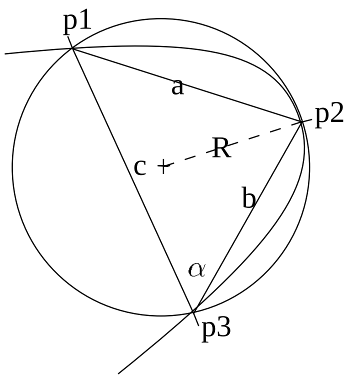


Figure 1: Curvature estimation

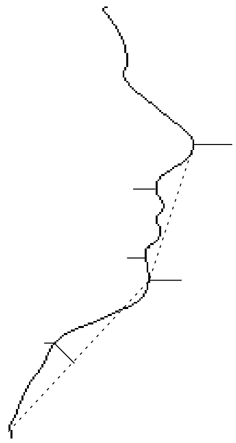

Figure 2: Features

# 7. FEATURE EXTRACTION

Many features relate to a distance measure. To be representative of the person, they must be normalised to be scale invariant, i. e. independent of the distance from the camera. The length used for normalisation is the distance separating the throat and the top reference points. This relatively large distance compensates for those point imprecisions, as demonstrated by practical results.

Two global features are extracted from the reference points: the mean width of the face around its middle height and the global curvature of back hair. Both are not very accurate, especially depending on hair, but they are sufficiently different among people to help classification.

The other features concern the frontside face outline, called the profile in what follows, and localised from the full outline thanks to the top and throat reference points.

The curvature of the profile is computed for each point as explained in the previous section. The nose and chin reference points are refined by picking the closest point with maximal curvature. By the way, the curvature amplitude and sign of the nose and chin help detecting face inconsistency.

As depicted in Fig. 6, the Chin-Nose section typically reveals seven maxima. The three median ones relating to mouth and lips are not always present and may vary largely. For this reason, they are not used and only four curvature maxima are used as features. Their signs are used as face hypothesis verification. The three distances separating those four maxima are the next features. As the two inner maxima (concavities) are less accurate, the sum of the three distances was added as feature, providing for a very accurate cue. The last feature concerning the Chin-Nose section is the area of the curve defined thanks to the Chin-Nose segment.

Three characteristics are extracted from the Chin-Throat part. As the throat could be limited by a collar, its associated reference point is not used. Instead a new point is defined, maximising the distance from the Chin-Throat segment. This distance is used as feature, as well as the curvature at this point and its distance to the chin (see Fig. 2).

Up to that point, 14 features are used as feature vector and face identification will take place thanks to the classification of those vectors, subject of the next section.

# 8. CLASSIFICATION

The aim of classification is to group faces based on the similarity of the feature vector extracted from them. Ideally, features extracted from the images of the same person should be very similar and the associated vectors should cluster in the N-dimensional space of features. In practice, some features undergo important variations due to the accuracy of the measure and the stability of the underlying face characteristics. The classification method must maximise the chance of assigning an image to its correct class. This depends on the quality of features (accuracy, relative discrimination power) and similarity of faces, and must thus be adapted to the problem and the features used. The ultimate aim is to have one person per class to achieve full identification. In case of doubt, the list of candidates is still a valuable information for search reduction or multiple sensor approach (fusion).

There are at least 4 approaches to perform classification. The first one relates to nearest neighbour techniques which assign a new object to the class which minimises some distance measure between the new object and the objects of the class. The second approach is to derive or estimate probabilities for the application of the Bayesian formula to get class belonging probabilities. The third approach uses artificial neural networks to learn from the database the features/class associations, hoping for them to generalise correctly for new feature vectors. The last is Fuzzy classification, for which combination rules are used to handle belonging descriptions of features to classes.

Until now, the work has been concentrated on feature extraction and classification is performed thanks to a simple nearest neighbour classifier. The feature vector extracted from a new image is compared to all feature vectors of the database thanks to a measure of similarity (city block distance, sum of the absolute differences). Those measures are grouped by class and averaged to give the class score. The class scores are ordered increasingly and the first class is proposed to represent the person identity. However, in order to reduce false alarms (such as the identification of a person who is not in the database) the proposition is accepted only if the score is below a given threshold. This threshold is still heuristical because it depends on the features and on the database, what is not stable during development. A second key to confidence, the immunity, is the ratio of two class scores: the first bad class score and the desired class score. If a class (person) has typical feature values, its score will be low and the score of the other classes will be high, resulting in a high immunity. An immunity more than 1.0 means a right classification.

To improve classification results, the relative importance of each feature is weighted by the inverse of its variance among all the database entries. This is intended to compensate for the inaccurate or erroneous feature values. Of course, in doing so, we decrease the importance of features that naturally vary much among different persons. This normalisation could be done relative to the variance of the class, but until now, as involving few images per class, this may not be representative. Harmon[23] has tried both approaches.

Up to now, some limitations affect the classification scheme. First, the relative importance of the different features is only controlled by variance, subject to caution. Secondly, the features are not independent what complicates the study of feature relevance but, depending on the classification scheme, the use of a combination of features can be valuable if more stable or discriminant. Third, the present scheme uses all the available features, even if some have unrealistic values. One solution is to limit the maximum possible distance for each feature.

As a conclusion, the present implementation is sufficient to have a complete route to identification. However, concerning its influence on the global identification efficiency, some new technique should be adopted. In that respect, knowing that features can be completely erroneous, a method looking for the relevant features is welcome.

# 9. RESULTS

The results are presented in the form of one simulation corresponding to the extraction of the 14 features, for the 81 images of the 20 persons.

Table 1 gives the rate of successful identification for different immunities, ratios of two error measures, respectively the best out-class match and the correct match. For instance, imposing an immunity of at least 1.1, 91% of the images were correctly identified. Depending on the confidence we need, we will impose some minimum level of immunity.

| Immunity | > 2.0 | > 1.5 | > 1.2 | > 1.1 | > 1.0 |
|---|---|---|---|---|---|
| Rate | 22% | 52% | 80% | 91% | 95% |

Table 1: Success rate versus immunity

A first result is that we always had the correct class in the first 2 best matches. We also had only 4 wrong matches, what corresponds to immunity values lower than 1.0. When we look at the class scores (not printed) of the wrong matches, we noticed their large values (above 10.0) in comparison with right matches (from 3 to 8). We thus have a means to reject our wrong matches (certainly with some good matches). Now, playing with such thresholds was not considered because other (adaptive) techniques can do the job right away.

Looking at the table, we see that 95% of the images are correctly identified with no security margin (immunity $> 1.0$). Practically, we must choose a minimum immunity greater than 1.0 to avoid false identification due to wrong matches. In the present case, the worst immunity was 0.84 (wrong match) so that we must set the minimum immunity to at least $1/0.84 = 1.2$.

There is still one important thing to mention before one can say the recognition ability reaches 80% for 20 persons. As you noticed, each person is represented by about four images. Since one of them is used as a test image, it remains only three in-class images (sometimes two). An error about one of these images (contour detection or feature extraction) has thus a large impact on the total rate: the intra-class score increases while the extra-class scores decrease and the immunity also decreases. So the rates will be more relevant when they will represent much more images, what will certainly increase the success rates.

# 10. PERSPECTIVES

As mentioned, the front view was first considered to achieve face identification. The major problem was the difficulty to find invariant features because extracting features was limited by the amount of measurable parts and invariance must consider illumination, rotation (3!), scale and body changes (hair, mouth and eyes may change a lot). On the contrary, the profile analysis has revealed easy, quick, robust and accurate localisation of chin and nose. Forehead and eye will certainly share the same properties thanks to new hypotheses (e.g. no glasses).

The natural next step is to combine both approaches, starting with the profile to get features and reference points and analysing the front view for new features and confirmation. As both views are rather independent (a human can hardly identify a profile view when used to the front view only), we hope to multiply the identification performance of both approaches (diminished by the presence of common features). Two perpendicular cameras can be used, what implies hardware (acquisition board) and software (registration) overheads. Instead, we will try to use a mirror at 45 degrees to record on the same image both views. The scale and vertical positioning will be consistent, reducing registration work. Mention that the profile view also gives accurately the distance of the head from the camera and its up/down orientation.

As all faces show the same configuration of the same parts, and thanks to the prior knowledge of reference points, model matching[19] seems appropriate. The questions are the availability of usable models and their possibility to offer parameters useful for classification. The simplest template to be considered is the eyebrow, mainly consisting of its length and curvature. The length is certainly a discriminant factor but the curvature, as modified by expression, may be less directly relevant.

Perhaps the most promising method consists in changing from sensor. Cameras offer a 2-D image of a 3-D object, resulting in rotation and scale incertitude. Their sensitivity to lighting conditions and the rather smoothness of face surfaces make the use of 3-D estimation from 2-D images unlikely or not accurate enough. On the other hand, a direct 3-D measure such as a range finder doesn't depend on ambient light, allows for background removal (based on distance) and 3-D transformation (rotation and scale). Absolute measures are available (e. g. nose length). Mapping a grey level image on a so-acquired 3-D description would be a very descriptive representation of a face.

Figure 3: 256x256 image from camera

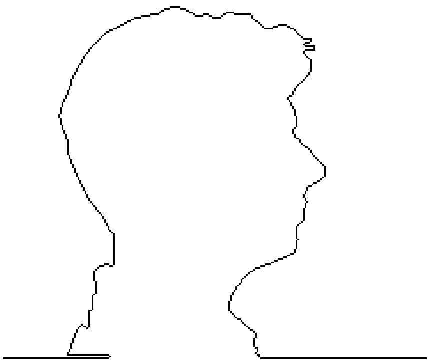

Figure 4: Head outline

## 11. REFERENCES


[1] R. Laughery and M. Wogalter, "Forensic Applications of facial memory research", In *Handbook of research on Face Processing*, Amsterdam, Netherlands: North-Holland, p. 544, 1989.

[2] A. Deffenbacher, "Forensic facial memory: time is of the essence", In *Handbook of research on Face Processing*, Amsterdam, Netherlands: North-Holland, p. 563, 1989.

[3] B. A. Golomb, D. T. Lawrence and T. J. Sejnowski, "Sexnet: A Neural Network Identifies Sex from Human Faces, *Advances in Neural Information Processing 3*, pp. 572-577, 1991.

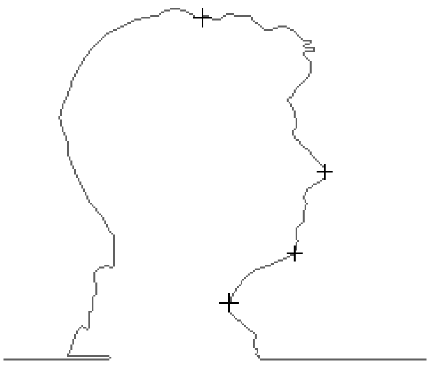

Figure 5: Reference points on outline

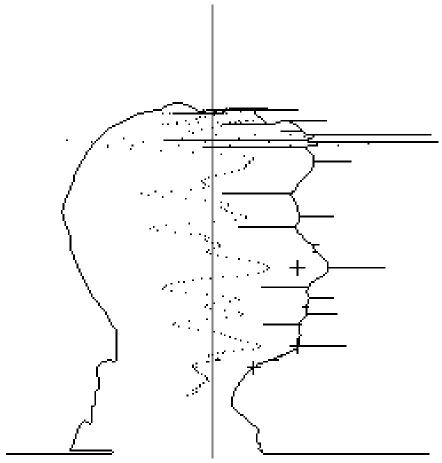

Figure 6: Outline curvature and its maxima


[4] G. W. Cottrell, "Extracting Features from Faces Using Compression Networks: Face, Identity, Emotion, and Gender Recognition using Holons", *Connectionist Models: Proceedings of the 1990 Summer School*, Morgan Kaufmann, San Mateo, pp. 328-337, 1990.

[5] N. M. Brooke, "Mouth Shapes and Speech", In *Processing Images of Faces*, V. Bruce and M. Burton (Eds.), Norwood, New Jersey: Ablex Publishing Corporation, pp. 20-40, 1991 .

[6] E. D. Petajan. "Automatic lipreading to enhance speech recognition", *Proceedings of the Global Telecommunications Conference*, pp. 265-272, Atlanta, GA: IEEE Communication Society, 1984.

[7] J. Sergent, "Structural processing of faces", In *Handbook of research on Face Processing*, Amsterdam, Netherlands: North-Holland, p57, 1989.

[8] J. L. Mundy, "Model-based vision, an operational reality ?", *Applications of Digital Image Processing XIV*, SPIE Vol. 1567 pp. 124-141, 1991.

[9] J. Blanc-Garin, "Faces and non-faces in prosopagnosic patients", In *Aspects of Face Processing*, Doordrecht, Netherlands: Martinus Nijhoff Publishers, p. 274, 1986.

[10] H. D. Ellis, "Introduction to aspects of face processing: Ten questions in need of answers", In *Aspects of Face Processing*, Doordrecht, Netherlands: Martinus Nijhoff Publishers, pp. 3-13, 1986.

[11] C. Perneel and M. Acheroy, "Face Recognition with Artificial Neural Networks", *IMACS Annals on Computing and Applied Mathematics*, Proceedings MIMS2 '90, Sept. 3-7, Brussels.

[12] I. Craw and D. Tock, "The Computer Understanding of Faces" In *Processing Images of Faces*, V. Bruce and M. Burton (Eds.), Norwood, New Jersey: Ablex Publishing Corporation, pp. 3-19, 1991.

[13] T. Kohonen, *Self-Organization and Associative Memory*, Springer series, 3th edition, 1989.

[14] T. J. Stonham (1986), "Practical face recognition and verification with WISARD", In *Aspects of Face Processing*, Doordrecht, Netherlands: Martinus Nijhoff Publishers, pp. 426-441, 1986.

[15] Huang Ho-Chao, Ouhyoung Ming and Ja-Ling Wu, "Automatic feature point extraction on a human face in model-based image coding", *Optical Engineering*, Vol. 32, No. 7 pp. 1571-1580, July 1993.

[16] T. Akimoto, Y. Suenaga and R. S. Wallace, "Automatic Creation of 3D Facial Models", *IEEE Computer Graphics & Applications*, pp. 16-22, Sept 1993.

[17] A. D. Linney, "The Use of 3-D Computer Graphics for the Simulation and Prediction of Facial Surgery", In *Processing Images of Faces*, V. Bruce and M. Burton (Eds.), Norwood, New Jersey: Ablex Publishing Corporation, pp. 149-178, 1991 .

[18] N. D. Duffy, "Animation using Images Samples", In *Processing Images of Faces*, V. Bruce and M. Burton (Eds.), Norwood, New Jersey: Ablex Publishing Corporation, pp. 179-201, 1991 .

[19] A. Yuille, D. Cohen and P. Hallinan, "Feature Extraction from Faces using Deformable Templates", *Proc. 1989 IEEE CS Conf. on Computer Vision and Pattern Recognition*, IEEE Computer Society Press, Los Alamitos, Calif., 1989, pp. 104-109, 1989.

[20] A. J. Goldstein, L. D. Harmon and A. B. Lesk, "Identification of human faces", *Proceedings of the IEEE*, 59, pp. 748-760, 1971.

[21] A. W. Young and H. D. Ellis (Eds.), *Handbook of research on Face Processing*, Amsterdam, Netherlands: North-Holland, 1989.

[22] D. M. Wuescher and K. L. Boyer, "Robust Contour Decomposition Using a Constant Curvature Criterion", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 1 pp. 41-51, 1991.

[23] L. D. Harmon, S. C. Kuo, P. F. Ramig and U. Raudkivi, "Identification of human face profiles by computer", *Pattern Recognition*, Vol 10, pp. 301-312, Pergamon Press Ltd. 1978.