Automatic Profile Identification

Charles Beumier, Marc Acheroy

Royal Military Academy
Signal & Image Centre (c/o ELEC)
Avenue de la Renaissance, 30 B1000 Brussels
e-mail beumier@elec.rma.ac.be

Abstract. This paper presents a practical implementation of a person verification system based on face profiles in a cooperative environment. The image acquisition constraints are given and motivated. A set of global and local features are automatically extracted from the profile outline of the head. The candidate profile is then compared with profiles of the database to issue a distance from which to accept or reject the claimed individual. To increase robustness, the temporal consistency of the decisions over several shots is examined, what is possible thanks to the overall speed of the developed system.

Keywords: person verification, profile, curvature.

1. Introduction

Computerized Face Recognition receives more and more interest as depicted by the ever growing number of papers [1] published in the field for a few years. Reasons such as security needs and computer improvements explain this phenomenon although face recognition has always intrigued brain researchers respectively to this famous ability of human beings.

Several approaches have been considered for computerized face recognition according to the exploited information (face, profile, 3D, static or dynamic). The majority of research programs considers static frontal or profile views. Frontal views offer many pieces of information (hair, eye, nose, mouth, chin, ...), many of which are subject to caution (hair, eye and mouth may vary in large proportions). On the other hand, profile views present obvious (in the sense of a computer program) reference points (nose, mouth, chin) and much information lies in the outline. For this reason among others (see next section), we considered person identification thanks to the outline of their profile, from hair to throat.

The two well known paradigms for face recognition are global matching and feature extraction followed by classification. Global matching algorithms consider the face or parts of it as full pieces to be matched against references. These algorithms are normally simple but require preprocessing to cope with translation, rotation, scale and light variation. Feature based methods need reproducible and discriminant characteristics to be extracted and compared with some database. The extraction procedure normally needs more code than global matching but provides for more control. Comparing and managing the database of references is also much quicker. These last advantages have led us to work with profile features.

In what follows, section 2 presents the objectives of the current development while section 3 indicates the working constraints. The two following sections detail image processing techniques employed for feature extraction of the profile. Section 6 explains how profile features are compared. Section 7 presents results of profile comparisons and section 8 introduces improvements due to a temporal analysis. Conclusions and references close the paper.

2. Objectives

This development was the integration of ideas considered during a research study carried out in 1993 and funded by MDN (see [2]). The conclusion of the study was that profile analysis is very attractive in regard to software complexity, execution speed and discrimination power (between individuals). The staticity of the head parts involved and the relatively good independency of the outline against head rotation increase robustness relative to attitude changes.

We kept the same kinds of acquisition constraints which simplify image analysis and provide for robustness (see next section). However, we modified the way the profile is analysed to let many (possibly redundant) features be considered in a rather simple program. We also took the opportunity to integrate both local (curvature) and global (angle) features.

But above all these minor changes, one new dimension has been exploited: time. Instead of specifying an acquisition time, we design the algorithm so that it can analyse several images per second, what lets it choose appropriate profiles and integrate decisions over time to improve robustness.

The main application of the current system is face verification for secure access where an individual claims his identity and waits for acceptance based on profile analysis. In the related european project (ACTS - M2VTS [3]), this information is to be combined with speech and face verification to enhance performance. The full system will be installed at several places for intensive evaluation in real conditions. A demonstrator, based on profile identification only, is running in our premises.

3. Framework

In the spirit of a cooperative environment, we adopted a number of person and scene constraints. First, the individual is asked to gaze in a general direction (to show his profile), to take his glasses and scarf off and to adopt a normal attitude. Secondly, only one person is allowed to stand in the field of view of the camera, in front of a uniform white background.

To increase resolution, the person sits on a chair, what minimizes height variation. We also rotate the camera 90 degrees to let the horizontal scanning (with more resolution) be parallel to the profile. If the contrast between the background and the head is not high enough, some extra lighting on the background is added.

The delivered images are typically 768x576, grey-level, with the profile at the top and the hair at the left. Only the front outline containing the forehead, nose, mouth, chin and possibly the throat and hair will be considered, typically providing for 500 useful contour points.

4. Contour Detection

As the profile analysis only considers the outline (for simplicity and quickness purposes), contour detection is crucial and has been tailored to provide for a simple and quick solution.

First, rough contour points are obtained as first large bright to dark transition from the top of the image towards the bottom. From each of these points, two contour following procedures are initiated in opposite directions. The large amount of contours compensates for problems such as looping, wrong contours (in hair or clothes) or early stop. Partial contours are then merged into one curve of at most 800 points.

The uniform white background is important for transition (contour following) and white is crucial to avoid loosing the edge in the eyebrow, mouth or for black people.

5. Profile Analysis

Many different approaches have been compared for profile analysis.

a) "Critical Points" Our previous study [2] led us to the extraction of a set of features mainly based on maxima of curvatures ("critical points") and their inter-distances. The comparison of curves was carried out thanks to configuration (distance), local (curvature) and global (angle) information, what could fairly discriminate between 20 persons. But the number of critical points is limited, a lot of information is not contained in them (slope and extent of forehead, throat) and curvature amplitude is sensitive to local deformation (lips).

b) "Euclidean distance" A completely different approach was to compute a distance between curves, trying to make them fit. First the nose was localized to solve the translation problem. Secondly, one curve was rotated to minimize its distance with the other curve in the forehead and nose regions which are rigid parts. If the two curves don't overlap sufficiently at that level, they are considered as different. Thirdly, a global measure was computed as the sum of the distances between corresponding pixels of the curves. The correpondence was such that for any pixel of one curve, only the minimal distance with a set of pixels of the other curve was retained. This allows for small scale variation and local distorsion. The results were fair but the computation time and the memory required to store the curve was juged too large.

c) "Global Transformation" The objective is to transform the original list of points in an array of values so that curve comparisons are made quicker and easier. We tried curvature and angle transformations. Curvature values are rich of information and are translation and rotation invariant. Unfortunately, they are sensitive to local distorsion (mouth, chin) and noise. We thus preferred to compute a mean orientation at each point (local angle). The comparison of curves was a simple difference after a normalization in indices (noses and scale must correspond) and in values (global offset because angles are not rotation invariant).

The adopted method is an optimized version of c). The idea is to keep local (curvature) and global (angle) values at many points to offer many (possibly redundant) features.

First (see Fig. 1), the nose is detected by considering rather long flat zones followed by large vertical transitions. The nostril is then localized as the inflexion point just at the right of the nose. It is a precisely located reference point, the nose being sometimes too circular to present a stable reference point. The eye is finally looked for as the first concave region at the left of the nose. It will be used as the second reference point for rotation and scale normalization. The chin has also been tried but is not a static part and can be too circular to offer a good reference point. The localization of the eye is critical (no glasses!) but allows for a quick normalization.

From the nostril point, 80 curvature and angle values (40 at each side of the nostril) are estimated at a regular interval (relative to the scale). Curvature has been estimated by [2]

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

(where $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ are the 3 vectors of the triangle defined by 3 points on the curve), which, multiplied by 1000, gives values typically in the range -100..100. Angle values (in radians) are multiplied by 100 and rounded to the lowest integer.

6. Profile Comparison

Two profiles are compared relative to their feature list to issue a distance measure according to

$$dist = \frac{meanA * devA * meanC * devC}{100 * cntA * cntC}$$

where A and C stand for Angle and Curvature respectively, $mean$ is the mean difference between corresponding values, dev is the standard deviation of the difference, cnt is the number of feature values considered.

$cntA$ and $cntC$ take into account the fact that not all the feature values are involved in the distance measure. First, values not present in both curves are suppressed. Secondly, the features which are too different from one curve to the other (separately for angle and curvature) are not used.

Classically, three possibilities may arise according to the distance value: acceptance (very low distance) rejection (high distance) and doubt. In case of doubt another source of knowledge is requested, possibly thanks to a human operator (or another modality). The choice of the thresholds, application dependent, is typically done by analysing the FRR/FAR curves (see section 7).

Experiments have shown that the distance computed here is sometimes largely affected by some little influence. In practice, we solved that problem by storing sufficient feature lists of each individual in the database. A large distance between two feature lists does not necessarily imply that these come from different persons, but a recognition is assessed if there are entries showing low distances. The influence of normalization, which allows a rather high sensibility of the distance measure, has been depicted by a partition of database entries for a same person. According to their attitude, a person will elicit the response of a specific set of his database entries. The database has to contain sufficient entries to cover typical attitudes (normalization) while keeping entries sufficiently discriminant to the person (a 3/4 view, because flat, tends to be similar to most people).

7. Static Results

The performance of the system can be evaluated according to the False Rejection Rate (FRR) and False Acceptance Rate (FAR) curves. The first curve depicts the rejection rate of an acceptable person while the second curve depicts the acceptance rate of an intruder, both relative to the decision threshold. If we must ensure that no intruder is accepted, we choose a low threshold.

Those curves have been established by computing the distance of all possible profile pairs from the database. For each database entry, the minimum of the distances to entries of the same person is added to the FRR histogram, while the minimum distances to all other persons (independently) fill the FAR histogram. At the end, histograms are cumulated to offer the FRR and FAR curves. Minimal values have been used to avoid the sensitivity problem outlined at the end of the last section. In practice, a score based on best entries of each person is used (see section 8).

Fig. 2 shows the FAR and FRR curves for the database consisting of 267 images of 41 persons collected at many different times. We see that the rates are encouraging relative to the simplicity of the method. The EER (Equal Error Rate: where the curves cross) is 3.7 %.

The values are largely influenced by the quality of the database. In the previous measure, an entry with no similar entry for the same person penalises the FRR. A second test (see FRR_10 on Fig. 2), not considering such 'isolated' entries, implied an EER of 2.7 %, only dropping 10 entries (which had a minimal distance of at least 10 for entries of the same person).

As depicted in Fig. 2, it is advantageous to choose a small threshold to get a small FAR. However, reliability is obtained at the price of a reduced number of clear identifications. In practice, we used 5 as the maximal distance which votes for recognition. The fact that we can work at two images per second (Pentium

Fig. 1. Profile Image

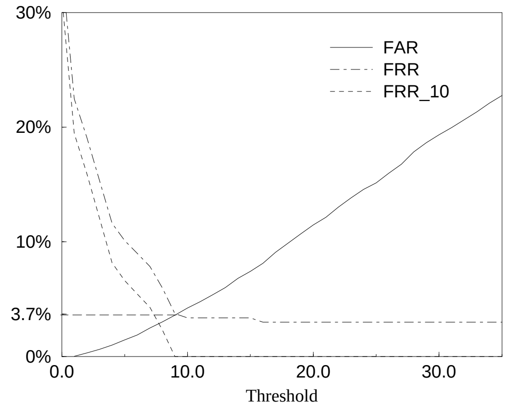

Fig. 2. FAR and FRR curves

75 Mhz) compensates for the possibly small number of identifications. On 41 persons, four couples of persons raised confusion, partially because they are still not well represented in the database but mainly because their profiles are similar.

8. Temporal Analysis

To release many human constraints (person's attitude, position), the system continuously acquires images. The algorithm allows to process several images per second and drops images which do not exhibit primary features such as a nose or eye. For the remaining images, only those with a small identification score are considered. As an advantage, it is no longer necessary to specify an acquisition time.

Time integration is not performed at the feature level. We preferred to compare each image separately and combine individual decisions. We indeed noticed some individual wrong identifications (for similar people) among correct ones. We thus delay decisions until we get sufficient consistent identifications. Here again, the rather high identification rate makes the system practical.

Globally, the system performance in real situations is summarised in Table 1, considering more than 2000 identifications. Each candidate image was compared to all database entries, yielding scores from 0 to 5 based on distances from 5 to 0 (section 6). Those scores were summed for each person and the one with the highest total score, if sufficiently higher than all other scores, was retained as the decision. We assess right to correct decisions, false to wrong decisions and doubt if no personal score is sufficiently higher than the others. Doubtful cases concern bad attitudes (large orientation, mouth largely opened), wrong normalization and (very few) bad contour detections. The rates have been obtained when people were asked to move, a human supervisor telling the clients when the identification was doubtful.

In the first row of Table 1, decisions were taken for each trial independently. A value of 29% of doubt is not so high considering that several images are analyzed each second. In the second row, eight consecutive decisions were combined: right identification if at least four correct individual decisions, wrong if at least four false individual decisions and doubt otherwise.

	Right	Doubt	False
Isolated decisions	69%	28%	3%
8-fused decisions	93%	6%	1%

Table 1. Global rates for continuous tests

In the long term, a temporal analysis is also necessary to keep the database uptodate. Individuals may be added or have a change (e.g. beard). The performance of the system relies heavily on the quality of the database. A minimal number of reference entries per person helps dealing with different attitudes but too many entries can increase wrong identifications and raise computation time. In practice, an average of 6 entries per person seemed sufficient, although atypical persons only needed 4 entries. New entries are added during registration or when an acceptable image is not correctly recognized. During identification, the system rewards rightly responding entries of the database and punishes wrongly responding ones. After sufficient presentations, each entry of the database has accumulated rewards and punishments on which to base the database cleaning or weighting. Largely punished entries should be removed because they participate to misidentification. If the database gets too large, we can also remove the entries with few rewards.

9. Conclusions

We have described an automatic profile identification system which has been designed to be simple, quick and practical. Based on normal acquisition constraints and dealing with the outline from hair to throat, it stores a small set of curvature and angle features. Comparing is quick and rather trivial.

The system gets its robustness from many facts. First, profiles rely mainly on rigid parts and are independent on small rotations. Secondly, the continuous identification scheme allows to base decisions on many comparisons and on different attitudes. Finally, the way the database is maintained allows for semi-automatic update. The weaknesses of the system, as agreed upon for the sake of simplicity and quickness, are the sensitivity to the eye detection and to long hair or beard.

Future research could consider more features (hair, ear, eye, ...) from the profile view to reach better performance or combine partial decisions more efficiently. But we prefer to analyse the facial surface as a new modality to offer new robust features.

10. Acknowledgements

This work has been supported by the M2VTS Project of the European ACTS program.

References

1. R. Chellappa, C.L. Wilson and S. Sirohey, "Human and Machine Recognition of Faces", In *Proceedings of the IEEE*, The institute of electrical and electronics engineers, inc., Vol. 83, No. 5, May 1995, pp. 705-740.
2. C. Beumier, M.P. Acheroy, "Automatic Face Identification", In *Applications of Digital Image Processing XVIII*, SPIE, vol. 2564, July 1995, pp. 311-323.
3. M.P. Acheroy, C. Beumier, J. Bigün, G. Chollet, B. Duc, S. Fischer, D. Genoud, P. Lockwood, G. Maitre, S. Pigeon, I. Pitas, K. Sobatta, L. Vandendorpe, "Multi-modal person verification tools using speech and images", In *Proceedings of the European Conference on Multimedia Applications, Services and Techniques (ECMAST '96)*, pp. 747-761, Louvain-La-Neuve, May 28-30 1996.
4. 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.
5. 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.