Human face profile recognition using attributed string

Yongsheng Gao\textsuperscript{a,\*}, Maylor K.H. Leung\textsuperscript{b}

\textsuperscript{a}Center for Graphics & Imaging Technology, Nanyang Technological University, Singapore 639798
\textsuperscript{b}School of Computer Engineering, Nanyang Technological University, Singapore

Received 27 July 1999; received in revised form 6 July 2000; accepted 1 December 2000

Abstract

A new attributed string matching method for human face profile recognition is proposed in this work. It is a novel idea to apply structural and syntactic technique on face profile matching. The approach works on a chain of profile line segments and highlights the favor of curve matching by suppressing the operations of insert and delete. The technique relies mainly on the merge and change operations of string to tackle the inconsistency problem of feature point detection. A quadratic penalty function is proposed to prohibit large angle changes and overmerging. The method produces very encouraging results and is found to be suitable for similar shape classification.

Keywords: Attributed string; Curve matching; Face recognition; Profile; Syntactic; Structural

1. Introduction

Face profile matching can be an important aspect for the recognition of faces. A face profile provides a complementary structure of the face that is not seen in the frontal view. The combination of the matching results of both the frontal and profile faces can improve the false acceptance rate. In addition, the system would be more foolproof because it is difficult to cheat the profile face identification by a mask. Profile analysis has also been used to assess the profile changes due to surgical correction \[1\] and create three-dimensional (3D) facial models \[2\].

Previous works on facial profile \[3–5\] extracted fiducial marks from the profile by heuristic rules, and a set of features was detected in terms of the positions of these fiducials. The distance between the features of a test profile’s fiducial points and that of the model’s fiducial points was calculated. The model in the database with the minimum distance was considered as the match. Kaufman et al. \[6\] reported a face profile recognition system using profile silhouettes. A set of normalized autocorrelations expressed in polar coordinates was used as a feature vector. The classification was based on a distance weighted $k$-nearest neighbor rule. The best performance of recognition reached 90%. Harmon et al. \[3\] manually drew the outlines from profile photos of 256 males. Nine fiducial points, i.e. the forehead, bridge, nose tip, nose bottom, upper lip, mouth, lower lip, chin and throat, were selected. A set of 11 features was derived from these fiducial points. After aligning the two profiles to be matched by two selected fiducial marks, the matching was achieved by measuring the Euclidean distance of the feature vectors derived from the outlines. They extended their work in Ref. \[7\] by reducing these 11 features to 10, because the nose protrusion was found to be highly correlated with other two features. Set partitioning technique was used to reduce the number of candidates to be included in the Euclidean distance measure. Decreased computation time was reported. In a continuous study \[4\], they defined 17 fiducial points, and thresholding windows were used to prune searching space. A 96% recognition rate was reported with or without the pruning technique. Wu et al. \[5\] developed a face profile recognition procedure based on 24 fiducial points. The differences from the work of Harmon et al. are as follows:
First, the outline curves were automatically obtained instead of by an artist’s drawing. Then, they used a B-spline to extract turning points on the outline curve. Subsequently, six interesting points and 24 features were derived from these points. A database of 18 oriental faces was used to test the performance of their approach. The stored features were obtained in a training process that used three profiles per person. Out of the 18 test images, 17 were reported correctly recognized. Yu et al. [8,9] proposed a rule-based fiducial mark extraction technique. Location and area constraints on nose lips and chin were used according to prior knowledge of human profile shape. Aibara et al. [10] proposed a method to recognize human face profiles based on P-Fourier descriptor (PFD). The P-Fourier descriptor is invariant to parallel translation and scale. They preprocessed the profile images by smoothing, edge detection, binarization, thinning and outline extraction to obtain outline curves of the profiles which consisted of 55 pixels upward from the nose tip and 90 pixels downward from the nose tip. A characteristic vector composed of 31 Fourier coefficients from the low frequency range was used. A 93.1% recognition rate was reported for 130 subjects in their experiment.

Most methods on profile recognition depend on the correct detection of fiducial points. Unfortunately, some features such as concave nose, protruding lips, flat chin, etc., make detection of such points difficult and unreliable. The human face profile is a highly structured geometric curve. From the viewpoint of representation, the set of fiducial points is a “sparse” representation of the underlying structures while the outline curve is a “dense” but honest representation of the shape. A high-level curve matching approach is, therefore, more appropriate and robust than point matching methods. A novel syntactic technique using attributed string is proposed here to recognize a chain of profile line segments rather than a set of inconsistent fiducial points. It highlights the favor of curve matching by suppressing the edit operations of “insert” and “delete”. The major operations are the “merge” and “change” of string primitives. A quadratic penalty function is proposed to prohibit large angle changes and overmerging. This technique provides strong discriminative power into the string matching method for similar shape classification and is found to be more accurate to distinguish one face from the other.

In the following, a brief overview of string matching together with its weakness on similar shape matching is highlighted in Section 2. In Section 3, the reason and principle of the merge dominant string matching approach are described in detail. An improved line attribute representation is adopted. Very encouraging experimental results are reported in Section 4. Finally, the paper is concluded in Section 5.

2. String matching

The structural and syntactic method is a high-level approach to find a symbolic and non-numeric description (e.g. string) of patterns [11,12]. It has been applied on shape recognition, character recognition, and speech processing. To the best of our knowledge, there has been no report on face recognition using string matching. The most related work using string is the process of two-dimensional (2D) shape, trademark and logo. With this approach, the symbolic representation of an input sample is matched against a number of models in a stored database. Symbols can be used as pattern primitives to represent 2D shapes [13–17]. A typical technique is to trace the outline of a shape using a series of primitives, each representing a different orientation. Then these symbols are transcribed into a string. One of the difficulties is the setup of suitable cost functions to compute the edit distance. Cortelazzo et al. [18] used string to code trademark contours and string match was applied for similarity measuring. Logo identification has been explored using algebraic and differential invariants [19] and string representation [14]. A hand-written character recognition method was reported by Nishida [20] that used structure to describe shapes.

For pattern recognition, the attributed string representation makes it easier to handle noise or distortion. It also has the advantage of smaller number of symbols to represent the shape of an object. Tsai and Yu [14] proposed an approach for shape recognition using attributed string with merging. They used attributed string to code line segments as the basic boundary primitives. Each line segment had two numerical attributes, i.e. its length and direction. However, it is found inadequate from our experiment to distinguish human face profile. The recognition rate is less than 20%. Face recognition using string matching is a challenging work. The differences among faces of different individuals are minor. Actually, all of them are within the class of “human face”. Therefore, particular attention should be paid on the discriminative power in the design of the classifier. In this work, enhancements on the attributed string representation capability and the cost functions of edit operations have been carefully designed. Special effort is made to tackle the problem of adding, missing and shifting of feature points. Subsequently, a novel merge dominant string matching process is proposed.

3. Merge dominant string match

The proposed string representation is based on the line segments generated from polygonal line fitting [21] on face profile outlines. Line segments are 2D entities with attributes of orientation, length, and the structural information of relative location with each other. The shape
of an object can be described as a set of ordered line segments using appropriate string representation.

3.1. Attributes determination

In order to enhance the representation power of attributes in Ref. [14] for similar shape matching, the line segment locations should be included. This can be achieved with the line primitive representation as \( P(l, \theta, x, y) \), where \( l, \theta, x \) and \( y \) are the length, orientation and midpoint location of the line, respectively. The line orientation \( \theta \) is defined as the minimum angle formed between the line segment and the reference line. The line between the nose tip and chin point is used as the reference line in this study.

3.2. String edit operations

The sequence of line segments from a face profile can be obtained through polygonal approximation or dominant points detection techniques. However, some objects, such as the face profiles shown in Fig. 1, lack sharp turning curvatures. This causes the adding, missing and shifting of feature points to occur. Conventional string matching method [14], that uses three types of edit operations (i.e. change, insertion and deletion) to transform one string into another, would behave badly under such circumstances. A merge dominant string matching method that suppresses insertion and deletion but encourages merging is proposed here to tackle this inconsistency problem.

Let \( A \) and \( B \) be two strings of line segment primitives, \( A(i) \) and \( B(j) \) be the \( i \)th and \( j \)th primitives in \( A \) and \( B \) with attributes \((l_i, \theta_i, x_i, y_i)\) and \((l_j, \theta_j, x_j, y_j)\), respectively. Define \( A(i:j) \) to be the sub-string from the \( i \)th to the \( j \)th primitives of \( A \), and \( A(i), B(j) \) to be \( A(1:i), B(1:j) \). The cost function for a change operation from \( A(i) \) to \( B(j) \), denoted as \( A(i) \rightarrow B(j) \), is defined as

\[
C(A(i) \rightarrow B(j)) = |l_i - l_j| + \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}
+ f(\Delta(\theta_i, \theta_j)),
\]

where

\[
\Delta(\theta_i, \theta_j) = \begin{cases} 
|\theta_i - \theta_j| & \text{if } |\theta_i - \theta_j| \leq 90^\circ, \\
180^\circ - |\theta_i - \theta_j| & \text{if } 90^\circ < |\theta_i - \theta_j| \leq 180^\circ, \\
|\theta_i - \theta_j| - 180^\circ & \text{if } 180^\circ < |\theta_i - \theta_j| \leq 270^\circ, \\
360^\circ - |\theta_i - \theta_j| & \text{if } 270^\circ < |\theta_i - \theta_j| \leq 360^\circ.
\end{cases}
\]

\( f() \) is a non-linear penalty function to map the angle to a scalar. It is desirable to ignore small angle variation but penalize heavily on large deviation. In this study,

![Fig. 1. Sample pairs of face profiles together with the detected dominant points.](image)

the function
\[ f(x) = \frac{x^2}{W} \]  

is used, where \( W \) is the weight to be determined experimentally.

A merge operation [14] with modified merge cost is used to tackle the problem of adding, missing and shifting of feature points. Let \( A(i-k+1:i) = P_{i-k+1}P_{i-k+2} \ldots P_i \) be a sequence of \( k \) primitives on a boundary to be merged, and \( A^k(i) \) be the merged primitive of these \( k \) primitives. One example is illustrated in Fig. 2 with \( k = 3 \). The merge operation is, therefore, denoted as \( A(i-k+1:i) \rightarrow A^k(i) \). For \( k = 1 \), it becomes \( A(i-k+1:i) = A(i:i) = A(i) \), which is the case without any merge operation. The cost function in merging \( k \) primitives is defined as

\[ C(A(i-k+1:i) \rightarrow A^k(i)) = f \left( \sum_{n=i-k+1}^{i} \Delta(\theta^k, \theta_n) \times l_n \right). \tag{4} \]

where \( r^k \) and \( \theta^k \) are the length and line orientation of the merged primitive \( A^k(i) \), \( l_n \) and \( \theta_n \) are the length and line orientation of primitive \( P_n \) before merging. Every \( \Delta(\theta^k, \theta_n) \) is weighted by the line segment’s length \( l_n \) because the contribution of angle difference of a primitive is assumed to be proportional to its length. The number of merge operations, \( k-1 \), is also taken into consideration of the cost. The cost function for a change operation after merge can be rewritten as

\[ C(A^k(i) \rightarrow B^j(j)) = |r^k - l_j| + \sqrt{(x^k - x^j)^2 + (y^k - y^j)^2} \]

\[ + f(\Delta(\theta^k, \theta_j)). \tag{5} \]

Eq. (1) is the special case of Eq. (5) with \( k = l = 1 \), which is the case without merge operation.

3.3. Merge dominant mechanism

For the costs of insert and delete operations, a null primitive \( A \) with zero length and indefinite angle and location is used. Suppose strings \( A \) and \( B \) are two profile strings from the same person. Intuitively, the new strings \( A^k \) and \( B^k \) after merging would have the same number of primitives with minimal merge costs. In addition, each corresponding primitive from \( A^k \) and \( B^k \) would resemble each other with little change cost. This would change completely with much higher costs if \( A \) and \( B \) are from different people. Thus, the matching process can be accomplished with just the merge and change operations. The insert and delete operations would only be useful if there are missing parts of the profile curve. Based on the above assumption, a merge dominant method is applied to encourage the merge operation but penalize heavily the insert and delete operations by increasing the costs of these two operations as

\[ C(A(i) \rightarrow A) = f(K_a) + l_i + K_{inc}, \tag{6} \]

\[ C(A \rightarrow B(j)) = f(K_a) + l_j + K_{inc}, \tag{7} \]

where \( l_i \) and \( l_j \) are the lengths of the \( i \)th and \( j \)th primitives in string \( A \) and \( B \), while \( K_a \) and \( K_{inc} \) are constants to represent the indefinite orientation and location of the line segment. For the purpose of penalization, the maximum angle difference, 90°, is assigned to \( K_a \), and the maximum location difference, the diagonal distance of the input image, is assigned to \( K_{inc} \). String matching is conducted according to Algorithm 1, where \( merge\_limitA \) and \( merge\_limitB \) are the controlled upper limits on the number of primitives to be merged into a new one in strings \( A \) and \( B \), respectively. More details of these numbers can be found in Section 3.4. \( D(i,j) \) is the edit distance or the minimum cost to match sub-string \( A(i) \) to sub-string \( B(j) \). Fig. 3 illustrates two examples in applying the technique to handle adding, missing and shifting feature points.

3.4. Merge number limit and computational complexity

The proposed technique is computationally expensive with time complexity as \( O(m^2 \times n^2) \) for strings of lengths \( m \) and \( n \). In order to cut down the computation time, an upper limit on the number (\( merge\_limitA = merge\_limitB \)) of primitives to be merged into a new one is imposed in this work. The computation complexity can then be reduced to \( O(m \times n \times merge\_limit^2) \). In this study, \( merge\_limit \) is found suitable to take a value \( \geq 3 \) for human face profile recognition (see next Section). Table 1 shows the average real computation time with and without merge number limit. The experiments were conducted on a SGI Octane workstation with 300 MHz CPU and 512 MB RAM. The average computation time was dramatically reduced from 5.00 to 0.07 s by adopting merge number limit of value 3. This result indicates that the speed of the proposed method is acceptable for real applications. However, it could be still computational demanding for large database compared with existing face profile matching techniques based on relative distances of fiducial points.

---

Fig. 2. An example of the merge operation.
4. Experimental results

A face database of 30 persons with two profile images per person from the University of Bern [25] was used to test the capability of the proposed approach. Each image is of the size 512 × 342 pixels with very high contrast. Some examples are shown in Fig. 4. The two profile images of each person were used as model and input, respectively. There are 60 matching experiments in total if the roles of model and input are interchanged. The nose tip and chin point of each profile face were detected automatically in the preprocessing stage. The line, L, between these two points was used to normalize image size, align face position and crop facial area to avoid processing subject’s hair. The face profile images were segmented by binarization, edge detection, thinning process to obtain 1 pixel width outline curves, facial area cropping, and feature point detection algorithm [21] to approximate the curves with line segments. Thus, every face profile outline was represented by a set of line segment end points. Samples of the profile pairs (models and test images) together with the detected dominant points have been shown in Fig. 1. Note that some pairs are very similar such that even human observers can hardly distinguish them. In addition, there are instances of adding, missing and shifting of feature points from one profile to its match.

To evaluate the overall recognition performance and analyze parameter sensitivity, two experiments were conducted with $W$ and merge limit as variables. $W$ was found easily tuned since the recognition rate remained higher than 96% when $W$ ranged from 30 to 100 (Fig. 5). Note that the recognition rate was as high as 100% when $W$ ranged from 40 to 50 with merge limit = 3.

It is found that the system without merging (i.e., merge limit = 1, which means only one line is merged into a line) could only correctly recognize 10% of the faces (Fig. 6). It improved quickly with merge operations and reached the optimal value of 100% when merge limit was 3. The system performance then degraded slightly to 98.33% and kept unchanged even when merge limit approached infinite. This result demonstrates that the system design is correct with excellent and stable performance.

In order to investigate the sensitivity of the proposed method to noise on profile curve, an experiment using the same strategy as in Ref. [24] was conducted. Random Gaussian noise with $\sigma_N$ ranging from 0.5 to 1.5 was added to the profile curve by rounding it off to the nearest integer and adding it to the digital curve. Examples of profile lines with different noise levels are illustrated in Fig. 7. The experimental results (Table 2) demonstrate that the method is robust to small noise. However, its performance degraded rapidly from 96.7 to 46.7% when $\sigma_N$ increased from 1.0 to 1.5.

It is interesting to compare the proposed method with other techniques as shown in Table 3. MHD is the modified Hausdorff distance [22] to match one template to the other in a point-to-point matching manner while LHD is the line segment Hausdorff distance [23] to match in a line-to-line fashion. The current method proposes a curve-to-curve matching approach. From Table 3, it is obvious that the results improve (become more accurate) from the point-to-point approach to the curve-to-curve technique. On the other hand, the former is
more flexible while the proposed might be restricted to well segmented curves. This is a reasonable trade-off between flexibility and accuracy.

Table 2
Recognition accuracies on profiles with different noise levels

<table>
<thead>
<tr>
<th>$\sigma_N$</th>
<th>0.5</th>
<th>1.0</th>
<th>1.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognition rate</td>
<td>100%</td>
<td>96.7%</td>
<td>46.7%</td>
</tr>
</tbody>
</table>

Table 3
Comparison of recognition rates from three different approaches (MDSM: merge dominant string matching)

<table>
<thead>
<tr>
<th>Method</th>
<th>MHD</th>
<th>LHD</th>
<th>MDSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognition rate</td>
<td>93.3%</td>
<td>96.7%</td>
<td>98.33%–100%</td>
</tr>
</tbody>
</table>
5. Conclusion

We have proposed, in this paper, a new attributed string matching method for human face profile recognition. The technique relies mainly on the merge and change operations of string to tackle the inconsistency problem of feature point detection. Hence, the proposed approach suppresses the string edit operations of insert and delete. A quadratic penalty function is used to prohibit large angle changes and overlapping. This makes the proposed technique robust to the errors from previous low-level processing that are usually inevitable in practice. The method produces very encouraging results and is found to be suitable for similar shape classification.

Compared with previous human face profile recognition methods, our method employs line segments derived objectively from the outlines instead of fiducial marks based on heuristic rules that may be invalid for unusual faces. Moreover, line segments are more reliable and accurate than fiducial marks to represent the distinctive details of an object. The experimental results also indicate that merge dominant string matching approach provides a new way to enhance the discriminative capability of string matching for the application of similar object classification.

The proposed method is a curve-to-curve similarity measuring. Every part of the whole curve is taken into consideration and contribute equally during the matching procedure. It is important for similar curve classification to make use of all distinctive information on curves. In contrast, this would cause one limitation of the method, that is, it cannot tackle the problem of partial matching on partially occluded faces. Future work would be on the sensitivity analysis of the proposed method to pose variation. Other directions could look into the possibility to fuse it with LHD to include the inner face features in the matching process.

References

About the Author—YONGSHENG GAO received the B.Sc and M.Sc degrees in Electronic Engineering from Zhejiang University, China, in 1985 and 1988, respectively, and his Ph.D. degree in Computer Engineering from Nanyang Technological University, Singapore. Currently, he is an assistant professor with Nanyang Technological University. His research interests include computer vision, pattern recognition and face recognition.

About the Author—MAYLOR K.H. LEUNG received the B.Sc degree in Physics from the National Taiwan University in 1979, and the B.Sc, M.Sc and Ph.D. degrees in Computer Science from the University of Saskatchewan, Canada, in 1983, 1985 and 1992, respectively. Currently, Dr. Leung is an associate professor with Nanyang Technological University, Singapore. His research interests include structural pattern processing, face recognition, motion analysis and string processing.