Face Recognition from Face Profile Using Dynamic Time Warping

Bir Bhanu and Xiaoli Zhou
Center for Research in Intelligent Systems
University of California, Riverside, CA 92521, USA
{bhanu, xzhou}@ vislab.ucr.edu

Abstract

Most of the current profile recognition algorithms depend on the correct detection of fiducial points and the determination of relationships among these fiducial points. Unfortunately, some features such as concave nose, protruding lips, flat chin, etc., make detection of such points difficult and unreliable. Also, the number and position of fiducial points vary when expression changes even for the same person. In this paper, a curvature-based matching approach is presented, which does not require the extraction of all the fiducial points, but uses information contained in the profile. The scale space filtering is used to smooth the profile and then the curvature of the filtered profile is computed. Using the curvature value, the fiducial points, such as nasion and throat can be reliably extracted using a fast and simple method. Then a dynamic time warping method is applied to match the face profile portion from nasion to throat based on the curvature value. Experiments are performed on two profile face image databases. Recognition rates and conclusion are presented and discussed.

1. Introduction

Face profile is an important aspect for the recognition of faces, which provides a complementary structure of the face that is not seen in the frontal view. Though it inherently contains less discriminating power than frontal images, it is relatively easy to analyze and more foolproof. Within the last decade, several algorithms have been proposed for automatic person identification using face profile images. Most of these algorithms depend on the correct detection of all fiducial points and the determination of relationships among these fiducial points.

Tangency-based techniques assume that by choosing the appropriate reference point, there will be a line through the reference point that is tangent to the face profile at one of the fiducials. Harmon et al. [12] use manually entered profile traces from photographs of 256 male faces. They locate eight independent fiducials on the profiles and obtain the ninth fiducial by rotating a point from the chin about the pronasale until it intersects the profile above the pronasale. Later, Harmon et al. [4] increase the number of fiducials from nine to eleven, and achieve 96% recognition accuracy for 112 subjects, using a 17-dimensional feature vector. The most significant problem with tangency-based techniques is that there is not a line that is bitangent to the pronasale and chin for profiles with protruding lips [9].

Curvature-based techniques overcome some of these limitations and are invariant to rotation, translation, and uniform scaling. Campos et al. [5] analyze the profile of the face using scale space techniques to extract eight fiducials. This technique assumes that there will be nine zero-crossings on the profile, and this assumption could be invalidated by facial hair particularly moustaches and the hairline on the forehead. Dariush et al. [6] extract nine fiducials based on the observation that the curvature of the profile alternates between convex and concave, with the point of maximal absolute curvature in each segment corresponding to a fiducial. Cartoux et al. [7] use face profiles extracted from 3D range images for face recognition. Gaussian curvature is used to extract the pronasale and nasion from a 3D range image. Gordon [8] use range data and curvature values for face recognition. The technique uses high level sets of relationships of depth and curvature to define various face descriptors.

Akimoto et al. [1] use a template matching approach to find the position of the same five fiducials used by Galton [3]. Their technique uses a template consisting of approximately 50 line segments to represent a generic face profile.

2. Technical Approach

2.1. Motivation

Although there are many methods for detecting fiducial points, there is still no single technique that can reliably extract all of the salient fiducials for every face. This is not surprising, because there are so many variations of a normal face profiles. The fiducials that researchers have extracted over the years have varied, although the five fiducials used by Galton in 1910 [3] are usually included in the set of fiducials. Some profiles are too difficult for all fiducials to be reliably extracted, so in these cases a feature vector approach based on the same fiducial points of different face profiles will fail.

Our research aims at overcoming the limitation of extracting the same fiducial points for different face profiles under different situations. Therefore, we presents a curvature-based matching approach, which does not focus on all fiducial point extraction, but attempted to use as much as information as a profile possesses. The scale space filtering is used to smooth the profile and then the curvature of the filtered profile is computed. Using the curvature value, the fiducial points, including the nasion and throat can be reliably extracted using a fast and simple method after pronasale is decided. Then a dynamic time warping method is applied to compare the face profile portion from nasion to throat based on the curvature value. The reason of choosing dynamic time warping as the matching method is that it is much more robust distance measure for time series than Euclidean distance, allowing similar shapes to match even if they are out of phase in the time axis [2]. Overall technical approach is shown in Figure 1.

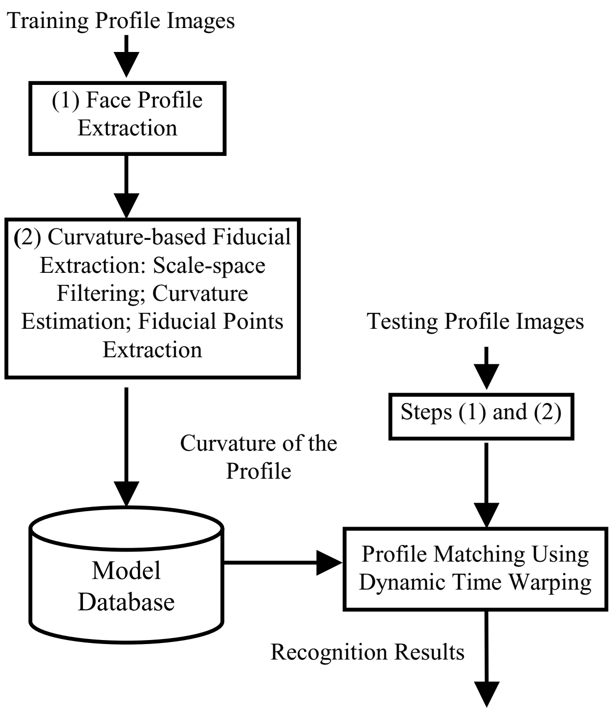

Figure 1. Technical Approach

2.2. Face profile extraction

The outline of a profile is treated as a function, which consists of a set of points $\Gamma(u) = (x_u, y_u)$, including fiducial points like nasion, pronasale, chin and throat. The following is the procedure, which is used to extract the face profile curve from the side view images.

- Apply canny edge detector to the grey-level image and obtain a binary image.
- Extract the outline curve of the front of the silhouette (the profile line) as the face profile contour by extracting the leftmost point different from background.
- The 2D curve can be regarded as 1D function $\Gamma(u)$, where x is a row index and y is a column index of a pixel inside a profile line.

2.3. Curvature-based fiducial extraction

- **Scale-space filtering:** In our method, Gaussian scale-space filtering is used to eliminate the spatial quantization noise introduced during the digitization process, as well as other types of high frequency noise. Another reason is to extract the fiducial points reliably. The Gaussian convolution of signal f(x) depends both on *x*, the signal's independent variable, and on σ, the Gaussian's standard deviation. The convolution is given by

$$F(x,\sigma) = f(x) \otimes g(x,\sigma) = \int_{-\infty}^{\infty} f(u) \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-u)^2}{2\sigma^2}} du, \quad (1)$$

where $\otimes$ denotes convolution with respect to *x*. The curve Γ is parameterized by the arc length parameter u:

$$\Gamma(u) = (x(u), y(u)). \quad (2)$$

An evolved version Γ_σ of Γ is then computed. Γ_σ is defined as

$$\Gamma_\sigma(u) = (X(u,\sigma), Y(u,\sigma)), \quad (3)$$

where

$$X(u,\sigma) = x(u) \otimes g(u,\sigma) \quad Y(u,\sigma) = y(u) \otimes g(u,\sigma).$$

- **Curvature estimation:** In order to find curvature zero-crossing points or extrema from evolved versions of the input curve, we need to compute curvature accurately and directly on an evolved version Γ_σ. Curvature κ on Γ_σ is given as:

$$\kappa(u,\sigma) = \frac{X_u(u,\sigma)Y_{uu}(u,\sigma) - X_{uu}(u,\sigma)Y_u(u,\sigma)}{(X_u(u,\sigma)^2 + Y_u(u,\sigma)^2)^{1.5}}, \quad (4)$$

where the first and second derivatives of *X* and *Y* can be computed from the following equations:

$$X_u(u,\sigma) = x(u) \otimes g_u(u,\sigma) \quad X_{uu}(u,\sigma) = x(u) \otimes g_{uu}(u,\sigma)$$

$$Y_u(u,\sigma) = y(u) \otimes g_u(u,\sigma) \quad Y_{uu}(u,\sigma) = y(u) \otimes g_{uu}(u,\sigma)$$

In our method, the absolute values of curvature are computed at a specific scale of edge contours. Then some of the local maxima of absolute curvature are chosen as corner candidates and are checked at lower scales of smoothing. The initial scale must be large enough to remove noise and small enough to retain the real corners. One of the advantages of the method is that it does not depend on too many parameters and it does not require any thresholds.

- **Fiducial points extraction:** Since the profiles include the hair and some other parts that are not reliable for matching, we need to extract the subpart of a profile from the total profile for effective matching. The method is based on the absolute values of curvature, which are computed at a specific scale of profiles. The points with the local maxima of absolute curvature at a specific scale are chosen as corner candidates and are checked at lower scales of smoothing.

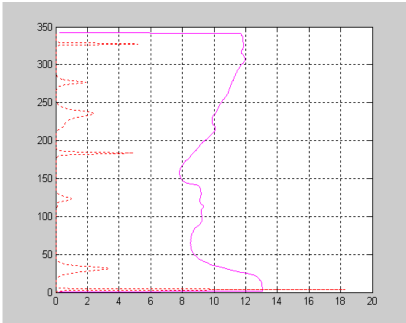

Figure 2. Face profile and curvature

We define pronasale as the leftmost point above throat in the middle part of the profile and nasion as the first point that has a large curvature value above pronasale. Finally we can get a portion of profile starting from nasion to throat. The smoothed profile and the absolute values of curvature are shown in Figure 2, and the portion of profile images for recognition is shown in Figure 3.

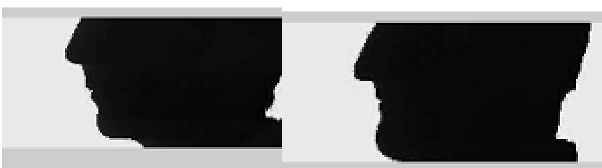

Figure 3. Face profile for comparison

The method of extracting the nasion and throat point is very fast and simple. It is described as follows:

- Find all points with the large curvature value as candidates and track them down to lower scales.
- Regard the rightmost point in the candidate set as the throat.
- In the profile image, regard the pronasale as one of the two leftmost candidate points in the middle part of the profile. Then check it using the curvature value around this point.
- Assume that there are no candidate points between pronasale and nasion, so the first candidate point above the pronasale is located as nasion.

The method can work well in our case since the images used only include the head above the neck. In other cases, the cropped images should first be obtained.

2.4. Profile matching using dynamic time warping

We choose the dynamic time warping as the matching method to compute the similarity of two profiles based on the absolute values of curvature, which are used to represent the shapes of face contours. The dynamic time warping is an algorithm to calculate the optimal score and to find the optimal alignment between two strings. This method is a much more robust distance measure for time series than Euclidean distance, allowing similar shapes to match even if they are out of phase in the time axis [2].

We use the Needleman-Wunsch [10] global alignment algorithm to find the optimum alignment of two sequences when considering their entire length. For two strings $s[1...n]$ and $t[1...m]$, we compute $D(i,j)$ for entire sequences, where i ranges from 1 to m and j ranges from 1 to n. $D(i,j)$ is defined as:

$$D(i,\ j) = min\{\ D[i-1,\ j-1] + d(s[j],\ t[i]),\ D[i-1,\ j] + gap,\ D[i,\ j-1] + gap\ \}. \quad (5)$$

Here, $d(s[j],\ t[i])$ represents the similarity between two points on face profiles. Since the face profile is represented by the curvature of the profile, $d(s[j],\ t[i])$ is calculated by Euclidean distance

$$d(s[j], t[i]) = \| s[j] - t[i] \| . \quad (6)$$

The penalty is defined for both horizontal and vertical gaps. It should be small and yet exist just to control non-diagonal moves. Generally, the penalties should be set to less than $1/10^{th}$ the maximum of the $d(s[j],\ t[i])$ [11]. In our method, we use the same constant penalty for both horizontal and vertical gaps. The maximum of $d(s[j],\ t[i])$ is about 5 and the gap penalties ranging from 0.5 to 1 perform well. The final score $D(m,\ n)$ is the best score for the alignment. Figure 4 gives an example.

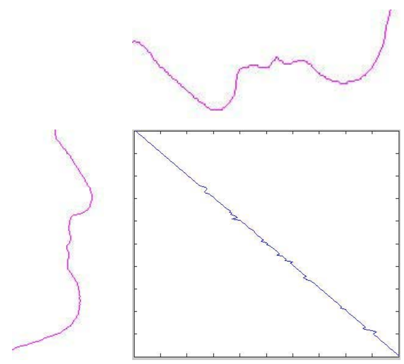

Figure 4. Warping of two face profiles (same person)

3. Experimental results

- **Data:** We first use face profile database from the University of Bern. It contains profile views of 30 people and three big grey-level profile images per person are with variations of the head position, the size and the contrast. The size of images is 342*512 pixels. In this experiment,

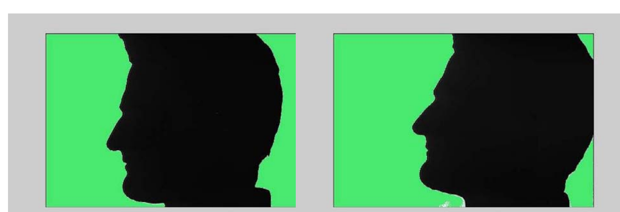

Figure 5. Face profile images from University of Bern

60 images with two images per person are used as the training dataset and the other 30 images are used as the testing datasets. Some examples of the dataset are shown in Figure 5.

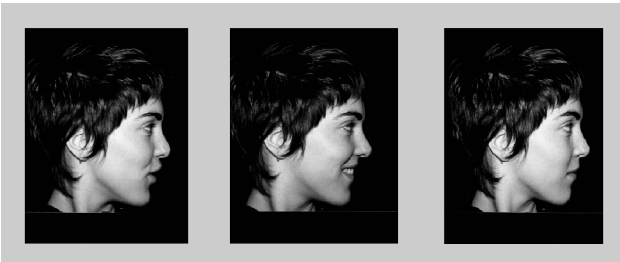

Figure 6. Face Profile Images from University of Stirling: Speaking, Smiling and Neutral Expression.

Another database we used to test our method is the set of face images from the University of Stirling, which contains 311 images of 35 people (18 females and 17 males). Thirty-one people (16 females and 15 males) have complete image set, which contains three poses (frontal view, view and profile view) and three expressions (neutral expression, smiling and speaking). The size of the image is 284*365 pixels. We use the images with neutral expression as the training data and the profile with smiling and speaking expression as the testing data for 31 people. Some examples of the dataset are shown in Figure 6.

- **Results:** The results for two databases are shown in Table 1. We can see that 90% profiles are correctly recognized for data from the University of Bern when the nearest neighbor method is used (3 errors out of 30 testing people). For the three persons who are not recognized correctly, if we choose the first three matching results as potential candidate, all of them are in the candidate list. So, in the future, we can use a refined verification process to improve the recognition rate.

For the database from the University of Stirling, the recognition rate is 78.4% for smiling face profiles and is 72.1% for speaking face profiles. Compared with the recognition rate of 90% for profiles from the University of Bern, although the results are degraded, the approach has potential since it can work without considering the number and position of fiducial points when the facial expression changes. This shows that our curvature-based matching method is relatively robust for face profile with different characteristics and under different situations.

<table>
<tr><td colspan="2">Data from University of Bern (30)</td><td>90.0%</td></tr>
<tr><td rowspan="2">Data from University of stirling (31)</td><td>Speaking</td><td>72.1%</td></tr>
<tr><td>Smiling</td><td>78.4%</td></tr>
</table>

Table 1. Recognition results

4. Conclusions

In this paper, a curvature-based matching approach is presented, which attempts to use as much as information the profile possessed. The scale space filtering is used to smooth the profile and then the curvature of the filtered profile is computed. Using the curvature value, the fiducial points, including the nasion and throat can be reliably extracted using a fast and simple way after pronasale is decided. Then a dynamic time warping method is applied to match the face profile portion from nasion to throat based on the curvature value.

Through the experiments, we can see that our curvature-based matching method is promising. Even for face profile with obvious expression variation, where the number and location of fiducial points are obviously different, the method can still work flexibly. Although the performance is degraded since the curvature is sensitive to noise to some extent, it can be improved by using a refined verification process in the future.

References

[1] T. Akimoto, Y. Suenaga, and R.S. Wallace. Automatic creation of 3D facial models. *IEEE Computer Graphics and Applications*, pages 16-22, September 1993.

[2] E. Keogh. Exact indexing of dynamic time warping. *Proceeding of the 28th VLDB Conference*, 2002.

[3] F. Galton, Numeralised profiles for classification and recognition. *Nature* 83, pages 127-130, Mar 31, 1910.

[4] L. D Harmon, M. K Khan, R. Lasch, and P. F. Ramig. Machine identification of human faces. *Pattern Recognition* 13, 97-110, 1981

[5] J.C. Campos, A.D. Linney, and J.P. Moss. The analysis of facial profiles using scale space techniques. *Proceedings of the IEE Colloquium on Machine Storage and Recognition of Faces*, pages.10/1-10/3, 1992.

[6] B. Dariush, S.B. Kang, and K. Waters. Spatiotemporal analysis of face profiles: Detection, segmentation, and registration. *Proceedings of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition*, pages 248-253, April 1998.

[7] J.Y. Cartoux, J.T. Lapreste, and M. Richetin. Face authentication or recognition by profile extraction from range images. *Proceedings of the IEEE Workshop on Interpretation of 3D Scenes*, pages 194-199, 1989.

[8] G.G. Gordon. Face recognition based on depth and curvature features. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* pp. 808-810, 1992.

[9] David O'Mara, Automated facial metrology. Ph.D. thesis. http://www.cs.uwa.edu.au/~davido/

[10] S.B. Needleman and C.D. Wunsch. A general method applicable to the search for similarities in the amino acid sequences of two proteins. *Journal of Molecular Biology,* pages 443-453, 1977.

[11] A. Gandhi. Content-based image retrieval plant species identification. http://cs.oregonstate.edu/~tgd/leaves/

[12] L. Harmon and W. Hunt. Automatic recognition of human face profiles. *Computer Graphic and Image Process*, vol.6, pages. 135-156, 1977.