A Scale-Space Approach to Face Recognition From Profiles

Zdravko Liposcak, and Sven Loncaric

University of Zagreb, Faculty of Electical Engineering and Computing
Department of Electronic Systems and Information Processing
Unska 3, Zagreb, Croatia
sven.loncaric@fer.hr

Abstract. A method for face recognition using profile images based on the scale-space filtering is presented in this paper. A grey-level image of profile is thresholded to produce a binary, black and white image, the black corresponding to face region. A pre-processing step then extracts the outline curve of the front portion of the silhouette that bounds the face image. From this curve, a set of twelve fiducial marks is automatically identified using scale-space filtering with varying the scale parameter. A set of twenty-one feature characteristics is derived from these fiducial marks. After normalising the feature characteristics using two selected fiducial marks, the Euclidean distance measure is used for measuring the similarity of the feature vectors derived from the outline profiles. Experiments were performed on a total of 150 profiles of thirty persons. Experimental results are presented and discussed. Finally, recognition rates and conclusions are given.

1. Introduction

Since the early 1990's Face Recognition Technology (FRT) become an active research area. A general statement of the problem of face recognition can be formulated as follows: Given still or video images of a scene, identify one or more persons in the scene using a stored database of faces [2].

The solution of the problem involves segmentation of faces from cluttered scenes, extraction of features from face region, identification, and matching.

Face recognition problems and techniques can be separated in two large groups: dynamic (video) and static (no video) matching. Dynamic matching is used when a video sequence is available. The video images tend to be of low quality, the background is very cluttered and often is more than one face present in the picture. However, since a video sequence is available, one could use motion as a strong cue for segmenting faces of moving persons.

Static matching uses images with typically reasonably controlled illumination, background, resolution, and distance between camera (or 3D scanner) and the person. Some of the images that arise in this group can be acquired from a video camera.

Machine recognition of faces has several applications, ranging from static matching of controlled photographs as in mug shots matching and credit card verification to surveillance video images. Mug shots matching is the most common application in static matching group. Typically, in mug shots photographs one frontal and one or more side views of a person's face are taken. Profile images provide a detailed structure of the face that is not seen in frontal images [7]. Face recognition from profiles concentrates on locating points of interest, called fiducial points. Recognition involves the determination of relationships among these fiducial points.

In this work we try to develop simple and fast method for detecting these fiducial points. For that purpose we treat the outline of a profile like a function and we use scale-space filtering [1], [3] for detection of extrema in that function and its first few derivatives. The profile is first expanded by convolution with Gaussian masks over a continuum of sizes. From this "scale-space" image we determine scale parameters which are used for detection of specific fiducial points. A set of twenty-one feature characteristics is derived from these fiducial marks. After normalising the feature characteristics using two selected fiducial marks, Euclidean distance measure was used for measuring the similarity of the feature vectors derived from the outline profiles [8]. Results of the proposed profile matching method in the presence of rotation, translation and size variance of profile faces are included in this paper.

2. Scale-Space Filtering

Scale-space filtering [1] is a method that describes signals qualitatively, in terms of extrema in the signal or its derivatives, in a manner that deals effectively with the problem of scale-precisely localising large-scale events, and effectively managing the ambiguity of descriptions at multiple scales, without introducing arbitrary thresholds or free parameters. The extrema in signal and its first few derivatives provide a useful general-purpose qualitative description for many kinds of signals.

Descriptions that depend on scale can be computed in many ways. As a primitive scale-parameterisation, the Gaussian convolution is attractive for a number of its properties, amounting to "well-behavedness": the Gaussian is symmetric and strictly decreasing about the mean, and therefore the weighting assigned to signal values decreases smoothly with distance. The Gaussian convolution behaves well near the limits of the scale parameter, σ, approaching the un-smoothed signal for small σ, and approaching the signal's mean for large σ. The Gaussian is also readily differentiated and integrated.

The Gaussian convolution of signal *f(x)* depends booth on *x*, the signal's independent variable, and on *σ*, the gaussian's standard deviation. The convolution is given by

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

where "*" denotes convolution with respect to *x*. This function defines a surface on the (x, σ)-plane, where each profile of constant σ is a Gaussian-smoothed version of f(x), the amount of smoothing increasing with σ. (x, σ)-plane is called scale space, and the function, F, defined in (1), the scale-space image of *f*.

At any value of σ the extrema in the *n*th derivative of the smoothed signal are given by the zero-crossings in the (n+1)th derivative. Although the scale-space filtering methods apply to zeros in any derivative, Witkin [1] restricted his attention to those in the second. These are extrema of slope, i.e. inflection points. In terms of the scale-space image, the inflections at *all* values of σ are the points that satisfy

$$F_{xx}= 0, F_{xxx}\neq 0, \qquad (2)$$

using subscript notation to indicate partial differentiation.

The contours of $F_{xx} = 0$ mark the appearance and motion of inflection points in the smoothed signal, and provide the raw material for a qualitative description over all scales, in terms of inflection points. Witkin applies two simplifying assumptions to these contours: (1) the *identity* assumption, that extrema observed at different scales, but lying on a common zero-contour in scale space, arise from a single underlying event, and (2) the *localisation* assumption, that the true location of an event giving rise to a zero-contour is the contour's x location as $\sigma \rightarrow 0$.

Referring to Fig. 1, notice that the zero contours form arches, closed above, but open below. The localisation assumption is motivated by the observation that linear smoothing has two effects: qualitative simplification- the removal of fine-scale features- and spatial distortion- dislocation, broadening and flattening of the features that survive. The latter undesirable effect may be overcome, by tracking coarse extrema to their fine scale locations. Thus, a coarse scale may be used to *identify* extrema, and a fine scale, to *localise* them. Each zero-contour therefore reduces to an (x, σ) pair, specifying its fine-scale location on the x-axis, and the coarsest scale at which the contour appears.

While coarse-to-fine tracking solves the problem of localising large-scale events, it does not solve the multi-scale integration problem. Witkin in [1] reduced the scale-space image to a simple tree, concisely but completely describing the qualitative structure of the signal over all scales of observation. In general, each undistinguished interval, observed in scale-space, is bounded on each side by the zero contours that define it, bounded above by the singular point at which it merges into an enclosing interval, and bounded below by the singular point at which it divides into sub-intervals.

3. Profile Face Analysis Using Scale-Space Filtering

The input in our face processing and analysis system is a gray-scale image of a scene containing the side view of a human face. The goal of our approach is to find a simple and fast method for recognition of profile faces. For that purpose we are using scale-space filtering and scale-space images of different faces for detecting the possible unique one or more parameters σ for all faces (see Fig.1.). The Gaussian convolution of a signal f(x), which represents profile line, with this σ parameter may be used to identify and localise extrema in f(x). These extrema represent interest points like a nose peak, nose bottom, mouth point, chin point etc. (see Fig.2.a).

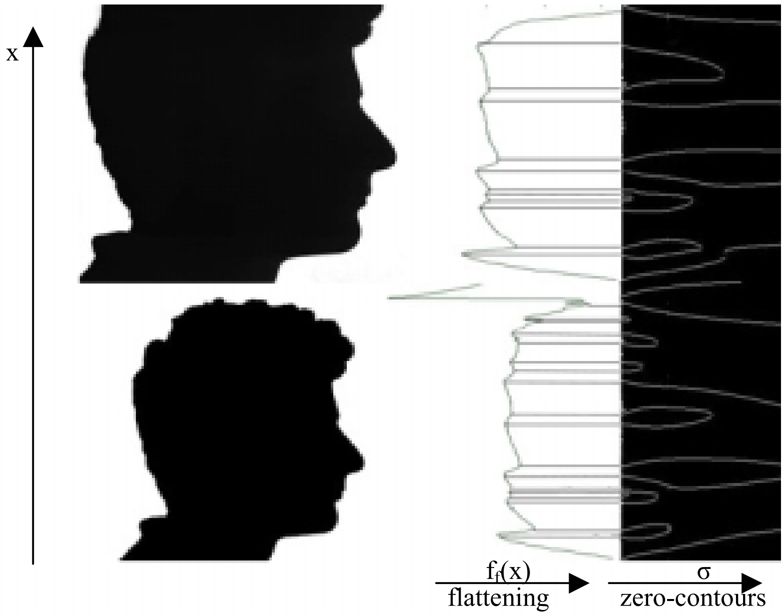

Fig. 1. Scale-space filtering of different profile face images (profile face images from University of Bern profile database[1])

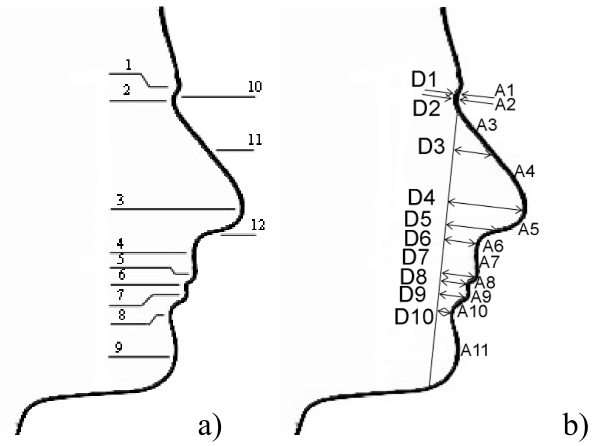

Fig. 2. a) The twelve fiducial points of interest for profile face recognition, b) Feature vector has twenty-one component; ten distances D_1- D_{10} (normalised with $/(D_4 + D_5)$) and eleven profile arcs A_1- A_{11}(normalised with $/(A_5 + A_6)$)

In first step the image was tresholded to produce a binary, black and white image, the black corresponding to the face region. A pre-processing step then extracts the outline curve of the front of the silhouette. Front silhouette line bounds the hair, face

---

[1] Copyright 1995 University of Bern, All Rights Reserved

and clothes of the person- this whole line we call the *profile line*. This profile line is converted from 2D picture into 1D signal $f(x)$ which we call the *profile vector*; where x is a row index and $f(x)$ is a column index of a pixel inside a profile line. For any row x without profile line we set $f(x) = 0$.

In the next step the profile vector f(x) is flattened using a Gaussian convolution $F(x,\sigma_L)$ of a profile with large parameter σ_L;

$$f_f(x) = f(x) - F(x, \sigma_L) \tag{3}$$

This step transforms all fiducial points in extrema and ensures better rotation invariance of the method (see Fig.1.).

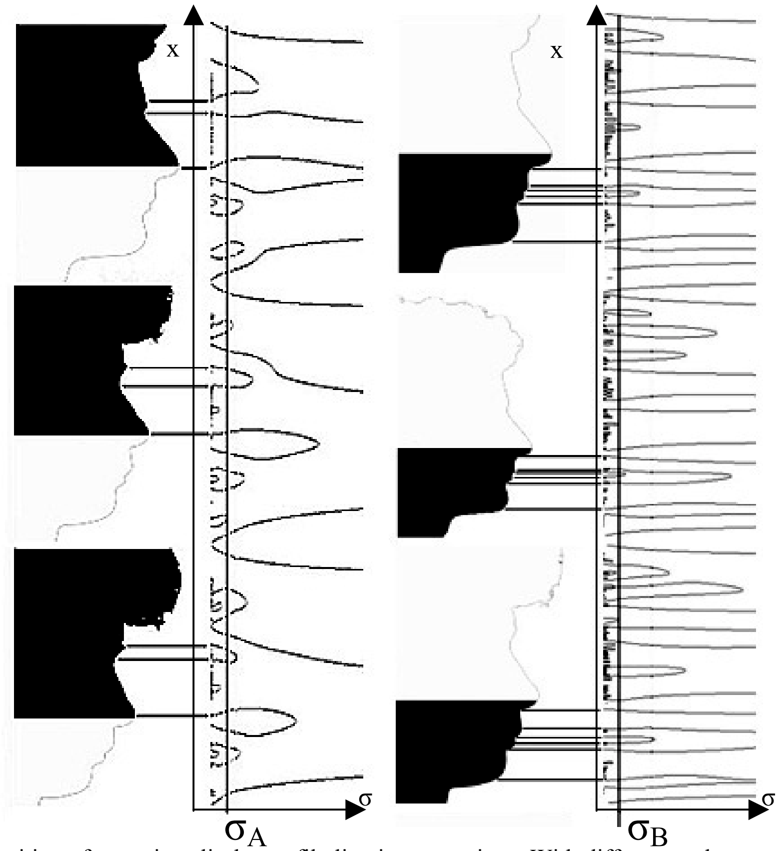

Fig. 3. Position of nose tip split the profile line in two regions. With different scale parameters σ_A and σ_B we detect all fiducial points.

The one-dimensional profile vector is expanded into a two-dimensional scale-space image. Comparing scale-space images of different profile faces we detect two unique parameters σ that gives extrema in desired fiducial points in all profile faces. After finding the fiducial points we can use the methods from similar papers [3],[4],[5] for describing profile faces using these points. From twelve fiducial points (see Fig.2.a; extrema points 1 to 9 and inflection points 10 to 12), we derive a simple set of twenty-one features. After normalising characteristics using two selected fiducial marks, Euclidean distance measure was used for measuring the similarity of the feature vectors derived from the outline profiles. A ranking of most similar faces is obtained by ordering the Euclidean distances.

We use two scale-space parameters σ - one for the profile region above the nose and one for the profile region below the nose. Two parameters are required because of different scales of facial features above and below of the nose tip.

Position of the nose tip splits the profile line in two regions; A – region above the nose and B – region below the nose. For this two regions we use different parameters σ_A and σ_B for detection of other points of interest (see Fig.3.). In region A and in region B of different profile images we can find unique parameters σ_A and σ_B for computing Gaussian convolution of flatten profile line and detection of global extrema that belong to the fiducial points 1- 2 and 4 – 9. For the inflection points 10 – 12 we can use the same σ_A and zero-contours of the inflection points.

Nose is an *event* that persists across large changes in scale. Next short algorithm is used for nose tip detection;

1. Extract the profile line f(x) from input graylevel picture.
2. Smooth the profile line using Gaussian convolution with small parameter σ_s (eq. 1.);
   $F(x, \sigma_s) = f(x) * g(x, \sigma_s)$
3. Smooth the profile line using Gaussian convolution with large parameter σ_L;
   $F(x, \sigma_L) = f(x) * g(x, \sigma_L)$
4. Flatten the profile line; $f_f(x, \sigma_s, \sigma_L) = F(x, \sigma_s) – F(x, \sigma_L)$
5. Compute Gaussian convolution with σ_N; $F_N(x, \sigma_N, \sigma_s, \sigma_L) = f_f(x, \sigma_s, \sigma_L) * g(x, \sigma_N)$
6. Find extrema in $F_N(x, \sigma_N, \sigma_s, \sigma_L)$.
7. Using simple rule $((\alpha_1 > 0)\ \&\ (\alpha_2 < 0)\ \&\ (d_1>d_2))$ for detection tip of nose we are find position of nose in all 150 pictures using only three unique parameters σ_s, σ_L and σ_N (see Fig.4.);

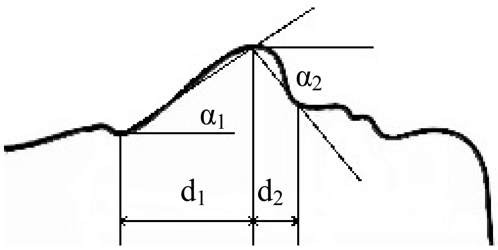

Fig. 4. Nose tip detection

The extreme in a smoothed and flattened profile that satisfy this rule belongs to the contour $F_x = 0$ of nose tip. Tracking the extreme by this contour we find position of nose tip. Coarse-to-fine tracking step is necessary because of large σ_L and σ_N.

The above algorithm with different values of σ_s, σ_L and σ_N is used for localising others fiducial points. Fiducial points are first two extrema above the nose tip and first six extrema below the nose tip and for this extrema we don't use coarse-to-fine tracking. All fiducial points are shown in Fig. 2a. From these points we derive the feature vector for each profile face in database. Feature vector has twenty-one component; ten distances D_1- D_{10} (normalised with $/(D_4 + D_5)$) and eleven profile arcs A_1- A_{11} (normalised with $/(A_5 + A_6)$) (see Fig. 2.b). D_1- D_{10} are distances between the *profile axis* and fiducial points. Profile axis is straight line, which pass through fiducial points one and nine.

4. Results and Discussion

Using the two face regions *A* and *B* with appropriate parameters σ we obtained the vector corresponding to each input profile face image. The dimensions of the profile vector depend on the number of fiducial points. In works based on different fiducial point extraction procedures [4], [5] authors use different numbers of fiducial points. From these fiducial marks they derived the sets of features. In [6] are defined 17 fiducial points which appears to be the best combination for face recognition. Most methods use the minimum Euclidean distance between the unknown and the reference feature vector to determine the correct identification of a profile, and some use thresholding windows for population reduction during the search for the reference feature vector. In our work we use twelve fiducial points; this number is detrmined by the nature of the method. Additional points can possible improve recognition rate and can be obtained by changing the scale parameters σ_s, σ_L and σ_N.

Table 1. Scale parameters for detecting fiducial points

Point	σ_s	σ_L	σ_N
nose tip (point 3)	6	17	28
points 1, 2 and 10, 11, 12	1	12	12.3
points 4, 5, 6, 7, 8 and 9	1	8	9

For training and testing set we used a profile face images Database University of Bern which contains profile views of the 30 people. For each person they took the five graylevel images with variations of the head position, the size and the contrast (1, 2, 3 big profiles with high contrast, looks like binarized, and 4, 5 small profiles with normal gray levels). Pictures are with controlled uniform background and without background clutter. Size of these images is 342 x 512 pixels.

Table 2. Feature vectors distance matrix (for the three persons)

Euclidean distances		Person 1			Person 2			Person 3		
		a	b	c	a	b	c	a	b	c
Person 1	a	0	0.16	0.53	0.56	0.61	0.62	1.08	0.99	0.98
	b	0.16	0	0.48	0.63	0.68	0.66	1.05	0.96	0.94
	c	0.53	0.48	0	0.84	0.95	0.85	0.66	0.56	0.54
Person 2	a	0.56	0.63	0.84	0	0.17	0.25	1.29	1.24	1.25
	b	0.61	0.68	0.95	0.17	0	0.29	1.41	1.37	1.37
	c	0.62	0.67	0.85	0.25	0.29	0	1.33	1.27	1.28
Person 3	a	1.08	1.05	0.66	1.29	1.41	1.33	0	0.17	0.25
	b	0.99	0.96	0.56	1.24	1.37	1.27	0.17	0	0.17
	c	0.98	0.93	0.54	1.25	1.37	1.27	0.25	0.17	0

Euclidean distances between feature vectors of nine profile pictures of three persons are shown in Table 2. We can see grouping of distances amount the three vectors that belongs to the same profile face.

In recognition experiment we use four profile images for each of 30 people in training set. The fifth profile face image for each of 30 person we used in testing set. In the first step we extract ten most similar profiles using only components $D_1 – D_{10}$. In the second step we used ten profiles from the first step and find the profile face with smallest Euclidean distance using component $A_1– A_{11}$. The algorithm is successively recognised 27 out of 30 persons. Profile faces from three people were recognised incorrectly. Obtained recognition rate is 90%.

5. Conclusion

In this paper, a new approach for face recognition from profile images is proposed. The method is based on scale-space analysis of the profile, followed by the extraction of fiducial points on the profile. Scale-space analysis gives the scale parameters for the family of profile faces needed for detection fiducial points. Using only nine constant scale parameters (see Table 1.) we find fiducial points in all 150 profile images. From these fiducial points a set of feature vectors is created. A ranking of most similar faces was obtained by ordering the Euclidean distances. Method is simple and fast and has shown promising results.

References:

1. A. P. Witkin: Scale-space filtering, Proc. 8th. Int. Joint Conf. AI 1983, pp 1019-1022.
2. R. Chellappa, C. L. Wilson, and S. Sirohey: Human and Machine Recognition of Faces: A Survey, Proc of the IEEE, Vol. 83, No. 5, May 1995. pp. 705-740.
3. S. Loncaric: A Survey of Shape Analysis Techniques, Patt. Rec. Vol. 31, No. 8, 1998, pp. 983- 1001
4. L. Harmon and W. Hunt, “Automatic recognition of human face profiles,” Computer Graphic and Image Process., vol.6, pp. 135-156, 1977.
5. C.Wu and J. Huang, "Human face profile recognition by computer," Patt. Recog, vol.23, pp. 255-259, 1990.
6. L. Harmon, M. Khan, R. Lasch, and P. Ramig, “Machine identification of human faces,” Patt. Recog., vol.13, pp.97-110, 1981.
7. Pigeon, S.: Multimodal Face Recognition and Identification. UCL. PhD Thesis. (1999)
8. Parker, J. R.: Algorithms for Image Processing and Computer Vision. Wiley Comp. Pub. (1998)