HUMAN IDENTIFICATION USING FACIAL CURVES WITH EXTENSIONS TO JOINT SHAPE-TEXTURE ANALYSIS

Chafik Samir, Mohamed Daoudi
GET/Telecom Lille 1, LIFL (UMR USTL-CNRS 8022) France

Anuj Srivastava
Department of Statistics, Florida State University, Tallahassee, FL 32306, USA

Keywords: Face recognition, Facial curves, Facial shapes, Geodesic of facial shapes.

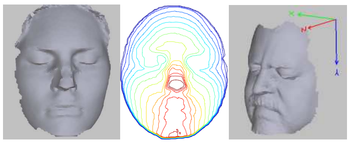
Abstract: Recognition of human beings using shapes of their full facial surfaces is a difficult problem. Our approach is to approximate a facial surface using a collection of (closed) facial curves, and to compare surfaces by comparing their corresponding curves. The method is further strengthened by the use of texture maps (video images) associated with these faces. Using the commonly used spectral representation of a texture image, i.e. filter images using Gabor filters and compute histograms as image representations, we can compare texture images by comparing their corresponding histograms using the chi-squared distance. A combination of shape and texture metrics provides a method to compare textured, facial surfaces, and we demonstrate its application in face recognition using 240 facial scans of 40 subjects.

1 INTRODUCTION

Automatic face recognition has been actively researched in recent years, and various techniques using ideas from 2D image analysis have been presented. Although a significant progress has been made, the task of automated, robust face recognition is still a distant goal. 2D Image-based methods are inherently limited by variability in imaging factors such as illumination and pose. An emerging solution is to use laser scanners for capturing surfaces of human faces, and use this data in performing face recognition (Chang et al., 2005), (Lu et al., 2006), (Bronstein et al., 2005). Such observations are relatively invariant to illumination and pose, although they do vary with facial expressions. As the technology for measuring facial surfaces becomes simpler and cheaper, the use of 3D facial scans will be increasingly prominent. A measurement of a facial surface contains information about its shape and texture (more precisely, the reflectivity function). In general, one should utilize both the pieces of information for recognition. Given 3D scans of facial surfaces and textured images, the goal now is to develop metrics and mechanisms for comparing their shapes and textures.

Our approach described in the paper (Samir et al., 2006) is to derive *approximate* representations of facial surfaces, and to impose metrics that compare shapes of these representations. We exploit the fact that curves can be parameterized canonically, using the arc-length parameter, and thus can be compared naturally. In addition to shapes of facial surfaces, we also consider the information associated with video images of the faces, referred to here as the texture images. The question is how to combine the shape information with the texture information to perform a joint face recognition. Motivated by a growing understanding of early human vision, a popular strategy for texture analysis has been to decompose images into their spectral components using a family of bandpass filters. Zhu et al. (Zhu et al., 2000) have shown that the marginal distributions of spectral components, obtained using a collection of filters, sufficiently characterize homogeneous textures. The choice of histograms as sufficient statistics implies that only the frequencies of occurrences of (pixel) values in the filtered images are relevant and the location information is discarded (Julesz, 1962). In addition to textures, these ideas have also been applied to appearance-based object recognition (Liu and Cheng, 2003). Following this approach, we will filter face images using Gabor filters (Gabor, 1946), and com-

Samir C., Daoudi M. and Srivastava A. (2007).
HUMAN IDENTIFICATION USING FACIAL CURVES WITH EXTENSIONS TO JOINT SHAPE-TEXTURE ANALYSIS.
In *Proceedings of the Second International Conference on Computer Vision Theory and Applications - IU/MTSV*, pages 253-256



(b)

Figure 1: (a): Examples of facial surfaces of a person under different facial expressions. (b) left: Examples of: a facial surface S, its corresponding facial curves $C_\lambda s$. (b) right: A coordinate system attached to another face.

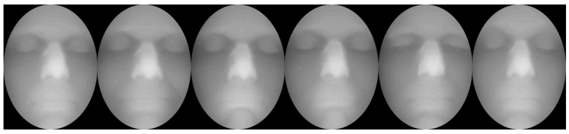


Figure 2: Top row: Range images of a subject's face under six different facial expressions. Bottom row: Range images of six different subjects under the same facial expression.

pare histograms of the filtered images using the χ^2-measure (Liu and Cheng, 2003). Rest of this paper is organized as follows: Section 2 describes a representation of a facial surface using a collection of facial curves, and presents metrics for comparing facial shapes under this representation. Section 3 presents a spectral representation of texture images using Gabor filters. Section 4 presents some experimental results and shows that a combination of shape and texture metrics improve the recognition rate. We finish the paper with a brief summary in Section 5.

2 REPRESENTATION OF FACIAL SHAPES

Let S be a facial surface denoting a scanned face. Although in practice S is a triangulated mesh with a collection of edges and vertices, we start the discussion by assuming that it is a continuous surface. Some pictorial examples of S are shown in Figure 1 (top row) where facial surfaces associated with six facial expressions of the same person are displayed. Let $F : S \mapsto \mathbb{R}$ be a continuous map on S. Let C_λ denote the level set of F, also called a **facial curve**, for the value $\lambda \in F(S)$, i.e. $C_\lambda = \{p \in S | F(p) = \lambda\} \subset S$. We can reconstruct S through these level curves according to $S = \cup_\lambda C_\lambda$. Figure 1 (a left) shows some examples of facial curves along with the corresponding surface S. In principle, the collection $\{C_\lambda | \lambda \in \mathbb{R}_+\}$ contains all the information about S and one should be able to analyze shape of S via shapes of C_λs. In practice, however, a finite sampling of λ restricts our knowledge to a coarse approximation of the shape of S. In this paper we choose F to be the **depth function**. Accordingly $F(p) = p_z$, the z-component of the point $p \in \mathbb{R}^3$.

2.1 Comparing Shapes of Facial Curves

Consider facial curves C_λ as closed, arc-length parameterized, planar curves. Coordinate function $\alpha(s)$ of C_λ relates to the direction function $\theta(s)$ according to $\dot{\alpha}(s) = e^{j\theta(s)}$, $j = \sqrt{-1}$. To make shapes invariant to planar rotation, restrict to angle functions such that, $\frac{1}{2\pi}\int_0^{2\pi}\theta(s)ds = \pi$. Also, for a closed curve, θ must satisfy the *closure condition*: $\int_0^{2\pi}\exp(j\,\theta(s))ds = 0$. Summarizing, one restricts to the set $\mathcal{C} = \{\theta | \frac{1}{2\pi}\int_0^{2\pi}\theta(s)ds = \pi, \int_0^{2\pi}e^{j\theta(s)}ds = 0\}$. To remove the re-parametrization group $\mathbb{S}^1$ (relating to different placements of origin, point with $s = 0$, on the same curve), define the quotient space $\mathcal{D} \equiv \mathcal{C}/\mathbb{S}^1$ as the shape space.
Let C^1_λ and C^2_λ be two facial curves associated with two different faces but at the same level λ. Let θ_1 and θ_2 be the angle functions associated with the these curves, respectively. Let $d(C^1_\lambda, C^2_\lambda)$ denote the length of geodesic connecting their representatives, θ_1 and θ_2, in the shape space $\mathcal{D}$. This distance is independent of rotation, translation, and scale of the facial surfaces in the $x - y$ plane. Now that we have defined a metric for comparing shapes of facial curves, it can be easily extended to compare shapes of facial surfaces. Assuming that $\{C^1_\lambda | \lambda \in \Lambda\}$ and $\{C^2_\lambda | \lambda \in \Lambda\}$ be the collections of facial curves associated with the two surfaces, two possible metrics between them are:

$$1 - \quad d_e(S^1, S^2) = \left(\sum_{\lambda\in\Lambda} d(C^1_\lambda, C^2_\lambda)^2\right)^{1/2}$$

$$2 - \quad d_g(S^1, S^2) = \left(\prod_{\lambda\in\Lambda} d(C^1_\lambda, C^2_\lambda)\right)^{1/|\Lambda|}.$$

d_e denotes the Euclidean length and d_g denotes the geometric mean. Here Λ is a finite set of values used in approximating a facial surface by facial curves.

The choice of Λ is also important in the resulting performance. Of course, the accuracy of d_e and d_g will improve with increase in the size of Λ, but the question is how to choose the elements of Λ. In this paper, we have sampled the range of depth values uniformly to obtain Λ.

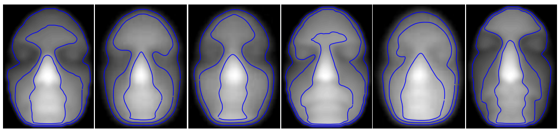
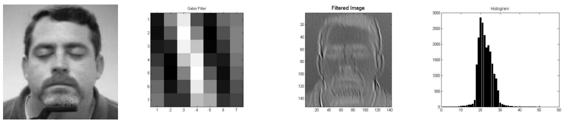


Figure 3: Three level sets in each surface. Top: same facial expressions, six different subjects. Bottom: six facial expressions, same subject.



(a) (b) (c) (d)

Figure 4: (a): Textured image (b): Gabor Filter (c): Filtered image (d): The spectral histogram of the image.

3 FACE RECOGNITION BY USING GABOR FACE REPRESENTATION

As mentioned earlier, we use a spectral decomposition approach to analyze and compare face images. The basic idea is to filter a given image using a collection of filters – Gabor filters, Laplacian of Gaussian, spacial derivatives, etc – and compute histograms of the filtered images to represent the original images. Furthermore, the original images can be compared by comparing their histograms using the χ^2 measure. The choice of filters in this approach is important and has a major bearing on the classification performance. In this paper, however, we restrict to a set of Gabor filters as they are most commonly used in the literature for appearance-based recognition. Consider a filtered image as a long vector and compute its histogram denoted by $f(\tilde{I})$. If $f_1(x)$ and $f_2(x)$ are two such histograms, perhaps generated from different images, then the (Pearson) chi-square statistic between them is:

$$\chi^2(f_1, f_2) = \int_x \frac{(f_1(x) - f_2(x))^2}{(f_1(x) + f_2(x))} dx \qquad (1)$$

This sets up the framework for texture based recognition. For any two images, I_1 and I_2, we define:

$$d_t(I_1, I_2) = \sum_{\alpha,\sigma} \chi^2(f(I_1 * F_{\alpha,\sigma}), f(I_2 * F_{\alpha,\sigma})) \ . \qquad (2)$$

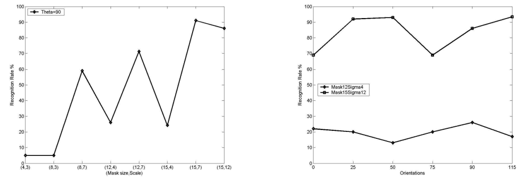
3.1 Performance for Different Scales and Orientations

Gabor filter is a frequency and orientation selective Gaussian envelope. The set of scale channels can be configured to capture a specific band of frequency components from an image. The set of the orientational channels are used to extract directional features. In order to determine which of the parameters of the filters is the most appropriate for our recognition task, we have changed the scales and the orientations of the filters and we compute the recognition rate for each parameter.

We fix the value of $\alpha = 90$ and we plot the recognition rate as a function of σ. The resulting curve is shown in Figure 5 (a). In figure 5(b) two curves representing the performance of two filters: the curve with squares represents the recognition rate against α for $\sigma = 4$ and the curve with diamond for $\sigma = 12$

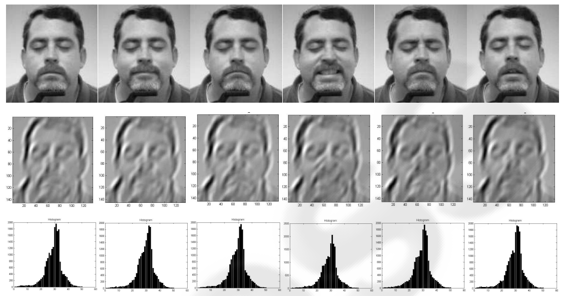
The first observation is that the performance of the recognition is affected by different combinations of scales and orientations. Therefore the best performance was achieved by taking the parameters settings of $\sigma = 12$, $\alpha = 115$.

Figure 6 shows an example of this filter applied on six images of the same person, filtered images and correspondent histograms are presented. The figure 6 shows the Gabor histograms obtained for six face expressions. These results show the robustness of the histograms to these expressions.



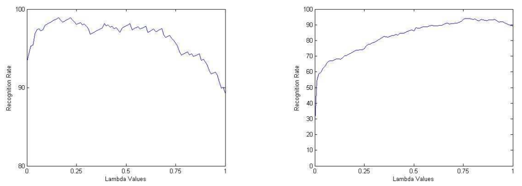
(a) (b)

Figure 5: (a): Recognition rate for $\alpha = 90$ against $\sigma = 4$ (b) The curve with squares represents the recognition rate against α for $\sigma = 4$ and the curve with diamond for $\sigma = 12$.
.



(a) (b) (c) (d) (e)

Figure 6: (a): Six facial expressions of the same person, their filtered images and their corresponding histograms.



(a) (b)

Figure 7: Recognition rate combination plotted versus λ(a): Recognition rate for $\sigma = 12$ and $\alpha = 115$ (b): Recognition rate for $\sigma = 4$ and $\alpha = 25$.

4 EXPERIMENTAL RESULTS

Each subject was scanned under six different facial expressions. Simultaneously a dataset of 2D color (texture) images was also collected for use in Gabor filter based recognition described later. In this section we present some experimental results to demonstrate effectiveness of our approach. As shown in the paper (Samir et al., 2006), the geometric mean d_g has given the best recognition rate results. A combination of shape and texture metrics provides a method to compare textured, facial surfaces is proposed. Indeed, in order to increase the accuracy of face recognition, it is often necessary to integrate the results obtained from different features of face: texture and shape.

Let d_t be the distance between two faces based on texture and d_g be the distance between two faces based on shape. Indeed, One of the difficulties involved in integrating different distance measures is the difference in the range of associated distances values. In order to have an efficient and robust integration scheme, we normalize the two distances values to be within the same range of $[0,1]$. The normalization is done as follows:

$$d_{tn} = \frac{d_t - d_t min}{d_{t\,max} - d_t min} \quad \text{and} \quad d_{gn} = \frac{d_g - d_g min}{d_g max - d_g min}$$

We define an integrated distance d between two faces as:

$$d = \lambda d_{tn} + (1-\lambda) d_{gn} \quad 0 \leq \lambda \leq 1 \tag{3}$$

The main problem is the choice of the value of λ. This idea is illustrated in Figure 7, where the recognition performance is plotted against λ. The results obtained in Figure 7 show that $\lambda = 0.17$ gives the best recognition rate 98.9.

5 SUMMARY

A new metric on shapes of facial surfaces was proposed in (Samir et al., 2006). In this paper, this method is extended to include analysis of facial textures in the recognition process. We use a spectral decomposition approach to analyze and compare 2D faces, the choice of filters in this approach is important and has a major bearing on the recognition performance. The experimental results clearly show that the combination of the texture information and the surface features of the same person outperform methods using one or other descriptor. For instance our method achieved a recognition rate of 98.9% in the case of recognizing faces under different facial expressions.

ACKNOWLEDGEMENTS

This work is supported by CNRS and GET under the project Recovis3D.

REFERENCES

Bronstein, A. M., Bronstein, M. M., and Kimmel, R. (2005). Three-dimensional face recognition. *International Journal of Computer Vision*, 64(1):5–30.

Chang, K. I., Bowyer, K. W., and Flynn, P. J. (2005). An evaluation of multimodal 2d+3d face biometrics. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(4):619–624.

Gabor, D. (1946). Theory of communications. *Journal of IEE (London)*, 93:429–457.

Julesz, B. (1962). A theory of preattentive texture discrimination based on first-order statistics of textons. *Biological Cybernetics*, 41:131–138.

Liu, X. and Cheng, L. (2003). Independent spectral representations of images for recognition. *Journal of Optical Society of America A*, 20(7):1271–1282.

Lu, X., Jain, A. K., and Colbry, D. (2006). Matching 2.5d face scans to 3d models. *IEEE Pattern Transactions on Pattern Analysis and Machine Intelligence*, 28(1):31–43.

Samir, C., Srivastava, A., and Daoudi, M. (2006). Three-dimensional face recognition using shapes of facial curves. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(11):1847–1857.

Zhu, S., Liu, X., and Wu, Y. (2000). Statistics matching and model pursuit by efficient mcmc. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 22:554–569.