

Delaunay 三角化於臉部辨識的應用

The Application of Delaunay Triangulation to Face Recognition

蔣依吾⁺、王瑞慶⁺
John Y. Chiang⁺, R. C. Wang⁺

中山大學應用數學系資訊組
⁺ Department of Applied Mathematics
National Sun Yat-Sen University
Kaohsiung, Taiwan 80424

張運龍⁺⁺
Yun-Lung Chang⁺⁺

中山科學研究院
⁺⁺ Chung-Shan Institute of Science and Technology
(CSIST) P. O. Box 90008-8-11, Lungtan
Taoyuan, Taiwan

摘要

這裡提出一種將臉部分割成三角形，不同於傳統方法的臉部辨識技術。Delaunay 三角化可根據三角形重心將圖形分割至符合要求，再將有相似期望值與變異數的三角形合併，最後根據這些圖形所畫出的長條圖作為辨認的判別式。Delaunay 三角化不會受到影像大小、旋轉及雜訊的影響。

關鍵字: 臉部辨識, Delaunay 三角化, 特徵擷取

Abstract

A new approach which differentiates from the conventional face recognition techniques in partitioning the facial image into triangular feature area is proposed. A Delaunay triangulation based on barycenters replaces triangles with successively smaller triangles until a stopping criterion is satisfied. A merging process combines adjacent triangles of similar means and variances. The histogram of the final set of areas is used as a discriminant to identify faces. The result of employing this Delaunay triangulation method is invariant to scale and orientation differences of the facial image tested, and robust to the contamination of noises.

Keywords: face recognition, Delaunay triangulation, feature extraction

1. Introduction

In this paper, a new approach which differentiates from the conventional face recognition techniques in partitioning the facial image into triangular feature area is proposed. A mesh of triangles is constructed throughout the image domain and the distribution of the

normalized triangular area is used as the primal feature for matching faces stored in the database. The triangular partitioning of the image plane has the advantage of being able to break away from the rigid horizontal and vertical segmentation schemes, better fit into the visual content and adapt to the natural edges of the image. Taking the distribution of the normalized triangular area as the similarity measure between the unclassified and the reference database images makes this recognition scheme invariant to scale change and orientation operations. The matching criterion can be further refined following the standard statistical methods to allow for inexact matching of the Delaunay triangles encountered in the noisy or moving environment.

Face recognition is a remarkable ability of human vision and has played an important role in the course of evolution. Different techniques have been proposed for the computer processing of face recognition. For feature-based techniques [1, 2], a set of geometrical feature vector is extracted from the facial views. Various statistical measures are then derived to describe the corresponding features and their relationship with regard to the facial topology. This approach is limited by the capability of selecting appropriate features that capture the sufficient information for the task given. The template-based methods apply a suitable set of metrics with templates to represent the whole face. Deformable template matching is a powerful technique for locating and describing object interested in an image when the structure of the object is previously known but the size, position, and orientation are not fixed [2, 3]. However,

each stored image template is always formed by picking a rectangular grid of points, instead of the triangular meshes discussed in this paper, as graph nodes. The level of deformability achievable is restricted by this rectangular lattice structure.

Triangular segmentations of the image plane have been proposed by Fisher [4]. The advantage of such a partitioning is to break away from the rigid right angle horizontal and vertical partitionings, better fit into the visual content and adapt to the natural edges of the image. The Delaunay triangulation is particularly well suited for adaptive meshing techniques [5], since it can be formulated as a sequential and local process. New points may be added and triangulated locally without the need for remeshing the domain in whole or in part.

The rest of the paper is organized as follows. The face recognition algorithm based on the Delaunay triangulation is described in Section 2. The proposed method has been implemented and tested on a database containing 15 frontal images of faces in the section followed. The chosen model images modified through different transformations, namely scaling, motion blur, rotation and noises-adding, are used as test inputs to demonstrate the feasibility of our recognition algorithm. The results show promising performance. Finally, conclusions are made and future works are discussed.

2. Face Recognition Algorithm

The focus of this paper is on developing a face recognition algorithm based on the Delaunay triangulation. The motivation for generating triangular meshes in a two-dimensional image plane lies in its ability to segment the face into natural edges better than the rectangular counterparts. Given a set of points in two dimensions, there exist many ways of joining them together to form a set of non-overlapping triangles. A Delaunay triangulation is a unique construction that no vertex from any triangle may lie within the circumcircle of any other triangle.

The face image is abstracted in terms of the Delaunay triangulation first. The distribution of the normalized triangular areas is then derived. The distribution of the unclassified image is compared by computing the squared difference, in turn, with those of the database images, returning a matching metric. The unknown person is then classified as the one giving the lowest difference value. This operation is invariant to the scale changes and orientation of an image since not the spatial coordinates, but the distribution of the normalized triangular area is used for squared difference computation. This strategy also allows for the inexact matching of the Delaunay graphs which may be caused by noise or motion blur [6].

The graph of the Delaunay triangulation can be defined by the triple $G = (V, E, F)$, where V is the set of vertices, $E \subset V \times V$ the set of edges and $F \subset V \times V \times V$ the set of the triangular faces. The triangular face can also be represented by a Cartesian triple of the node labels, i.e. $F \subset V \times V \times V$ such that $(i, j, k) \in F \Leftrightarrow ((i, j) \in E) \wedge ((j, k) \in E) \wedge ((k, i) \in E)$. The Delaunay triangulation of a vertex set V can be defined as the unique triangulation with "empty circles," i.e. no vertex lies inside the circumscribing circle of any Delaunay triangle, as follows:

$$DEL(V) = \{ (p_i, p_j, p_k) \in V^3 \mid C(p_i, p_j, p_k) \cap V - (p_i, p_j, p_k) = \emptyset \},$$

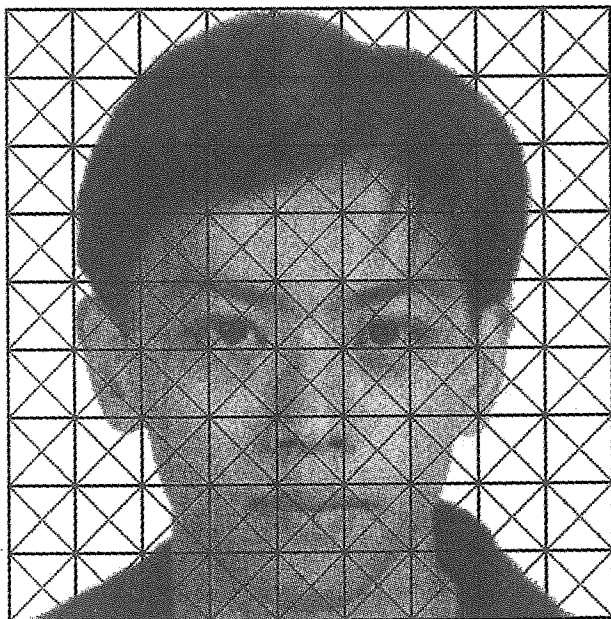
where $C(p_i, p_j, p_k)$ is the circle circumscribed by the three point p_i, p_j, p_k forming a Delaunay triangle.

Bowyer's algorithm employing the above circumcircle property is used to construct the Delaunay triangulation in a sequential manner [7]. It has been shown that Bowyer's algorithm exhibits linear computational complexity with the number of mesh points. Our method is based on a split-and-merge approach initialized on a regular grid of vertices as shown in Fig. 1 (a). A triangle is said to be homogeneous if the variance computed from its constituent pixel values is less than a prefixed threshold, i.e.,

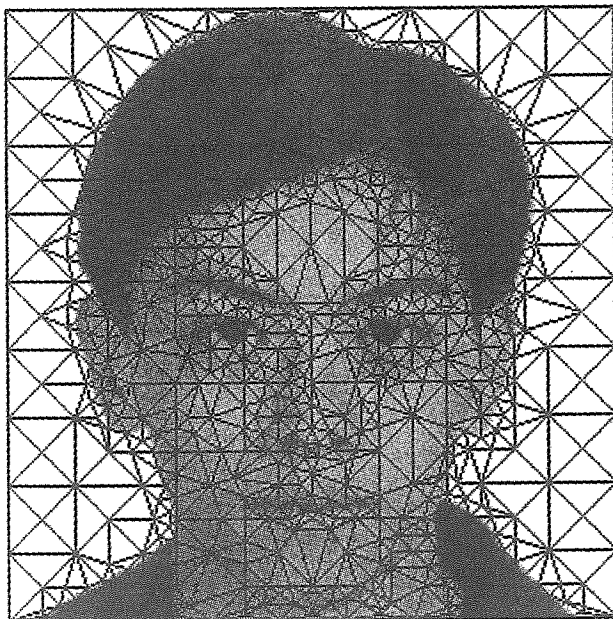
Triangle $f_i \in F$ is homogeneous

$$\Leftrightarrow V = \sum_{x,y \in f_i} \frac{(P_{x,y} - \bar{M})^2}{N_i} \leq T, \text{ where } \bar{M} = \sum_{x,y \in f_i} \frac{P_{x,y}}{N_i},$$

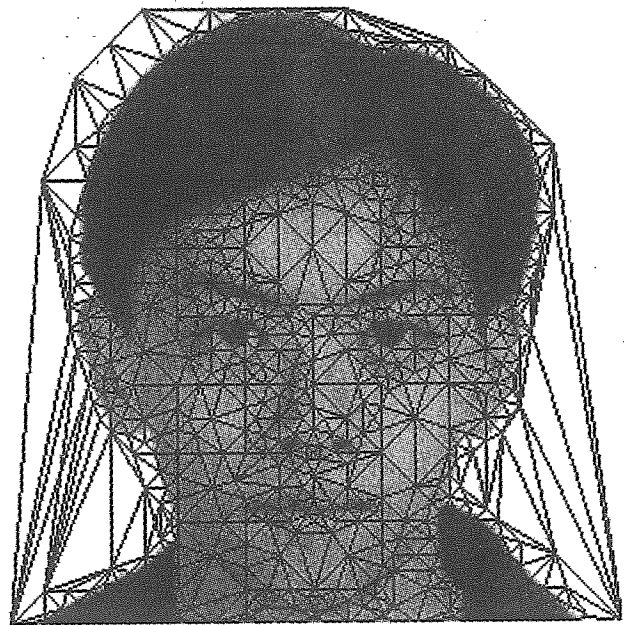
\bar{M} and V are the mean and variance of the triangle f_i , respectively. $P_{x,y}$ is the gray level of the constituent pixel (x, y) within the triangle f_i . T is a preset threshold and N_i the total number of the constituent pixels in the triangle f_i .



(a)



(b)



(c)



(d)

Fig. 1 (a) Initialization of the split-and-merge approach through a uniform grid structure. (b) An instance of the triangulation in the splitting process. (c) The final result after the merge step. (d) The histogram of the normalized triangular area distribution for the model image in part (a).

The splitting process progressively partitions the plane into finer triangular regions and converges when all the triangles are homogeneous. Each nonhomogeneous triangle is further split by the insertion of a point on its barycenter. The splitting of a larger triangle into three adjacent smaller ones is corresponding to a region full of delicate features (ref. Fig. 1(b)). The merge process is then followed to examine the condition that all the triangles sharing a vertex are similar with respect to their gray level mean and variance. These vertices are extracted and the relevant triangles merged. The merge step groups

adjacent triangles sharing common image characteristics together to reduce the complexity of the triangular mesh, as shown in Fig. 1 (c). The converged area of each triangle is computed and normalized with respect to the whole image size as follows:

$$n_{f_i} = \text{Area of the triangle } f_i / \text{Total area of image } I.$$

The number of the Delaunay triangles having the same normalized area n_{f_i} in image I is then tabulated and represented as $C_{n_{f_i}}^I$. This measure $C_{n_{f_i}}^I$ is used later in calculating the matching metric between the unclassified image and the model images in the database. The histogram of the normalized area distribution for a model image chosen is shown in Fig. 1(d). The distribution of these normalized triangular areas is used as the signature associated with a specific facial image. This normalization step makes the feature matching process independent of the scale of the unclassified image.

Since only the distribution of the triangular area, rather than the orientation of the triangular meshes, is considered, the recognition result is not affected by the rotation of the image tested. Some small perturbations in terms of the image gray-level changes require only local, rather than global, mesh re-structure. Therefore, this approach is robust to the sporadic noise contamination and motion blur.

3. Experimental Results

The database used throughout this research consists of 15 frontal images of faces. Persons participating sit down in front of a uniform background. The stored gray-scale images have a resolution of 256 x 256 pixels and a depth of 8 bits. These model images are triangulated first, then the corresponding normalized triangular area distributions are recorded and used for squared difference computation in later recognition stage. Fig. 2 shows a graph representing the ensemble

$(I, n_{f_i}, C_{n_{f_i}}^I)$ of the normalized area distribution for the database images, where I represents a specific image among the 15 model images contained in the database, n_{f_i} the normalized triangular area, and $C_{n_{f_i}}^I$ the corresponding count of instances for a normalized area n_{f_i} in the image I . For a fixed I value, the cross section of this graph corresponds to the image I 's distribution of the normalized triangular area. The recognition process proceeds by computing the squared difference between the distributions of the unclassified image and each of the facial images in the database. The unclassified image is declared as the model image with the lowest squared difference value.

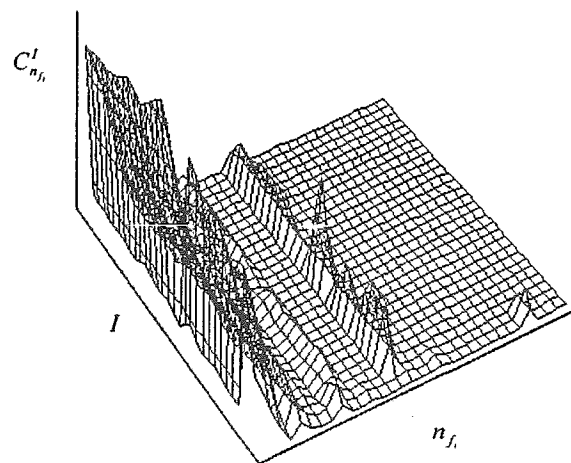


Fig. 2 The ensemble of the distribution of the normalized triangular area for the database images, where I is the image number in the database, n_{f_i} the normalized triangular area, and $C_{n_{f_i}}^I$ the count of instances for a normalized triangular area.

A person from our database is chosen for the purpose of illustration, as shown in Fig. 1 (a). In order to demonstrate the power of our method, the image is first transformed by motion blur for moving an amount of 10 pixels towards the lower right-handed corner, adding Gaussian noises in the facial area with a S/N ratio of 25 dB, 90 degree clockwise rotation and a 45% scale down, respectively, before proceeding the split-

and-merge triangulation stage. The images obtained after the above transformation are then partitions through the triangulation processing. The Delaunay triangulations for these unclassified test images are shown in Figs. 3 (a)-(d) and the corresponding normalized triangular area distributions in Figs. 4 (a)-(d). In comparison with Fig. 2, the rotation and scale-down versions of the unclassified images yield results of perfect matches. The distributions for the motion blur and noise-adding cases deviate from that of the model image slightly due to the local re-meshing property of the Delaunay triangulation. We then perform the squared difference calculation between the unclassified image A and each of the images B in the database as follows:

$$E = \sum_i (C_{n_{fi}}^A - C_{n_{fi}}^B)^2.$$

The squared difference E of the distribution of the normalized rectangular area distribution has a minimum value of zero that occurs if and only if two images match exactly. By using this squared difference test, the correct correspondence between the test image and the model image is reached in both cases.

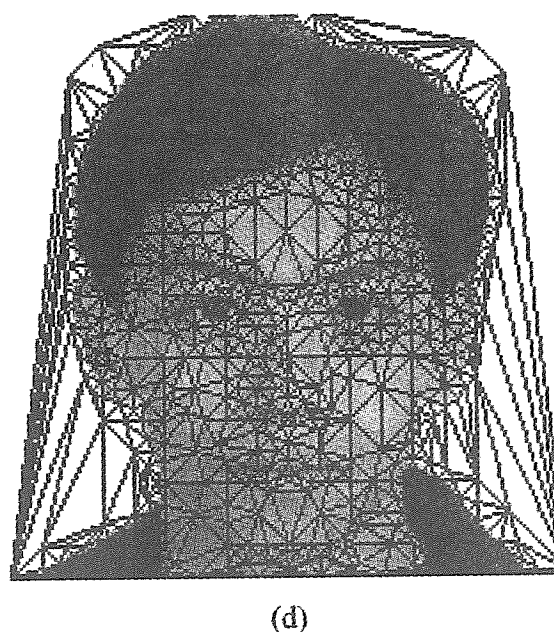
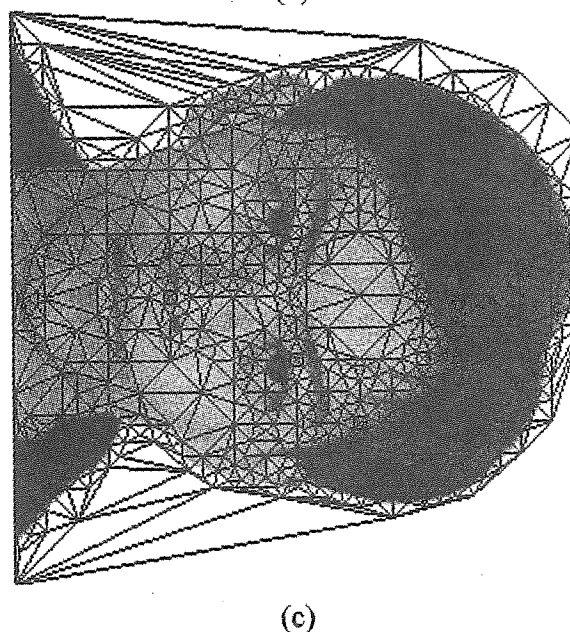
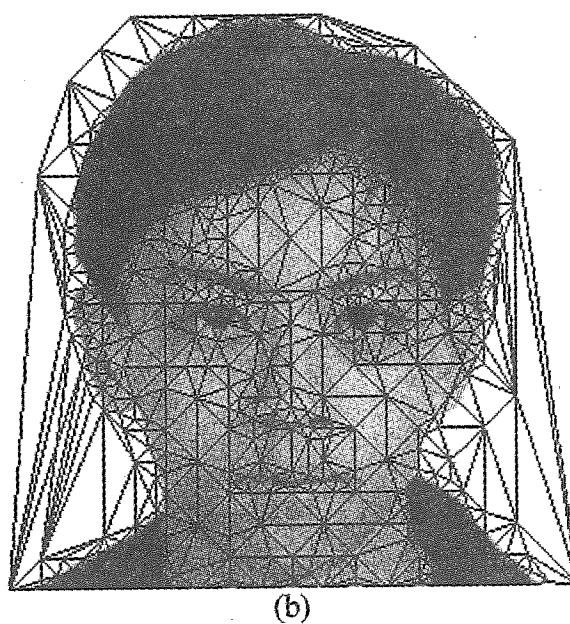
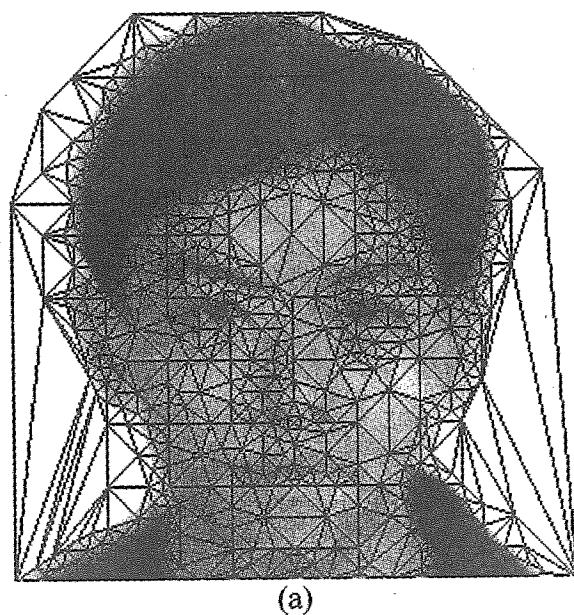


Fig.3 The Delaunay triangulation for an image suffered (a) motion blur for an amount of 10 pixels in the southwestern direction, (b) additive 25dB Gaussian noise in the facial area, (c) 90 degree clockwise rotation, and (d) 45% scale down.

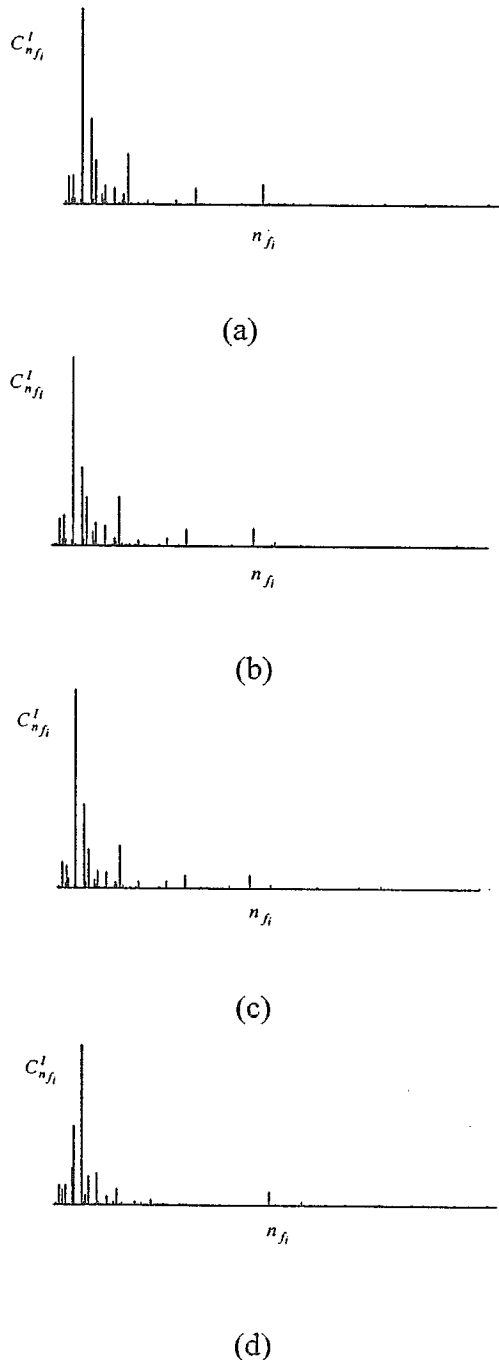


Fig. 4 The distributions of the normalized triangular area corresponding to the test images in Fig. 3. They are arranged in the order from top to bottom: (a) motion blur, (b) additive Gaussian noise, (c) rotation, (d) scale down versions, respectively.

4. Conclusions

A new face recognition algorithm based on Delaunay triangulation is discussed. The performance of this method is investigated using a database of 15 frontal facial images. The introduction of a lattice structure allows the facial features closely matched by the size of the triangular area. The distribution of the

normalized area is employed as the similarity measure in the squared difference test to compare the unclassified facial image with the model images in the database. The triangulation method primarily targets the problem of face recognition, but shall be general enough to be applied to other real world recognition tasks.

References

- [1] R. Brunelli and T. Poggio (1993), "Face Recognition: Features versus Templates," *IEEE Trans. Patt. Anal. Machine Intell.*, Vol. 15, No. 10, 1042-1052.
- [2] R. Chellappa, C. L. Wilson, and S. Sirohey (1995), "Human and Machine Recognition of Faces: A Survey," *Proc. IEEE*, Vol. 83, No. 5, 705-740.
- [3] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. v.d. Malsburg, R. P. Wurtz, and W. Konen (1993), "Distortion Invariant Object Recognition in the Dynamic Link Architecture," *IEEE Trans. Computers.*, Vol. 42, No. 3, 300-311.
- [4] Y. Fisher, Ed. (1995), *Fractal Image Compression - Theory and Application*, New York: Springer-Verlag.
- [5] D. G. Holmes and S. H. Lamson (1986), "Adaptive Triangular Meshes for Compressible Flow Solutions," in *Proceedings, International Conference on Numerical Grid Generation in Computational Fluid Dynamics, Karlsruhe, West Germany, 1986*, edited by J. Hauser and C. Taylor Pineridge, Swansea, UK.
- [6] A. M. Finch, R. C. Wilson, and E. R. Hancock (1997), "Matching Delaunay Graphs," *Pattern Recognition*, Vol. 30, No. 1, 123-140.
- [7] A. Bowyer (1981), "Computing Dirichlet tessellations," *The Computer Journal*, Vol. 24, No. 2, 162-172.