JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 22, 1047-1057 (2006)

New Approach for Static Gesture Recognition*

CHIN-CHEN CHANG, JIANN-JONE CHEN[1], WEN-KAI TAI[2] AND CHIN-CHUAN HAN
Department of Computer Science and Information Engineering
National United University
Miaoli, 360 Taiwan
[1]*Department of Electrical Engineering*
National Taiwan University of Science and Technology
Taipei, 106 Taiwan
[2]*Department of Computer Science and Information Engineering*
National Dong Hwa University
Hualien, 974 Taiwan

Static gestures can convey certain meanings and act as specific transitions in dynamic gestures. Therefore, recognizing static gestures is one of the most important aspects of gesture recognition. In this paper, a new approach is presented for recognizing static gestures based on Zernike moments (ZMs) and pseudo-Zernike moments (PZMs). The binary hand silhouette is first accommodated with a minimum bounding circle (MBC). The binary hand silhouette is then decomposed into the finger part and the palm part by morphological operations according to the radius of the MBC. After that, the ZMs & PZMs of the finger part and the palm part with different importance, respectively, are computed based on the center of the MBC. Finally, 1-nearest neighbor techniques are used to perform feature matching between an input feature vector and stored feature vectors for static gesture identification. Results show that the proposed approach performs better than previous methods based on conventional ZMs & PZMs in recognizing static gestures. The proposed technique could be useful in improving the recognition rate of static gestures.

Keywords: human computer interface, gesture recognition, static gesture recognition, Zernike moments, pseudo-Zernike moments

1. INTRODUCTION

The use of gestures in human computer interaction (HCI) has been extensively developed [2, 3, 6, 11-13, 17, 19]. Generally, gestures can be classified into static gestures and dynamic gestures. Static gestures are usually described in terms of hand shapes, and dynamic gestures are generally described according to hand movements. Since static gestures can convey certain meanings and act as specific transitions in dynamic gestures, recognizing static gestures is one of the most important parts in gesture recognition. Reliable descriptors to characterize static gestures are critical.

A number of researchers have developed algorithms to extract features for static gestures. Banarse and Duller [1] developed a neural network to perform static gesture recognition and used 2D plane cells as features for hand posture recognition. However,

Received August 16, 2005; accepted January 17, 2006.
Communicated by Jhing-Fa Wang, Pau-Choo Chung and Mark Billinghurst.
* A preliminary version of this paper has been presented at International Workshop on Advanced Image Technology 2002 (IWAIT 2002), Hualien, Taiwan, R.O.C., 2002.

the cost of extracting entire plane cells would be too high. Su *et al*. [15] presented a static hand gesture recognition system based on the flex angles of the ten fingers. It is effective and attractive but the control flowchart is too complicated. In addition, the training phase is needed to get the classification rule. Wu and Huang [18] presented a view-dependent hand posture recognition approach and used Fourier descriptors [20] to represent hand shapes. However, direct extension from 1D to 2D FFT is easy, but it would be inaccurate for translation invariance. Lee and Chung [10] proposed an algorithm that extracts features to recognize sign language based on orientation histograms of hand postures. However, the features of the same hand shapes would be affected by rotation. Zhou *et al*. [21] presented a bottom-up algorithm for static hand gesture recognition based on local orientation histogram features. This approach has advantages in terms of higher recognition accuracy and faster speed. However, the local orientation histogram features of static gestures would be affected by rotation. Schlenzig *et al*. [14] proposed a hand gesture interpretation based on the Zernike moments (ZMs) [8, 9, 16] of static hand shapes. Hunter *et al*. [7] proposed an algorithm for posture estimation based on the ZMs [7] of the hand silhouette for gesture recognition. However, ZMs characterize hand shapes globally and are therefore not applicable for local estimation.

Although these studies have provided many techniques for extracting hand shape features, there are still many problems in practical application. Further improvements are often required to extract more reliable hand shape features. In this paper, a new approach for static gesture recognition based on ZMs and pseudo-Zernike moments (PZMs) is proposed. Using the center of a minimum bounding circle (MBC) for translation invariance, more robust static gesture features can be obtained. The proposed approach can extract weighted ZMs & PZMs of static gestures, thereby describing static gestures globally with local support and providing a reliable description of static gesture feature. Results show that the proposed approach performed more accurately than previous methods based on conventional ZMs & PZMs. The proposed technique may significantly improve the recognition of static gestures.

Fig. 1 shows a diagram of the proposed approach. First, the binary hand silhouette is accommodated with an MBC. Then, the binary hand silhouette is decomposed into the

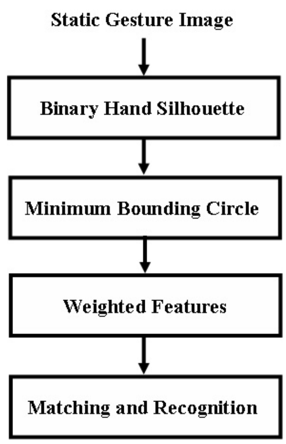


Fig. 1. Diagram of the proposed approach.

finger part and the palm part by morphological operations according to the radius of the MBC. After that, compute the ZMs & PZMs of the finger part and the palm part with different weights, respectively, based on the center of the MBC. Finally, use 1-nearest neighbor techniques [5] to perform feature matching between an input feature vector and stored feature vectors for static gesture identification.

The rest of this paper is organized as follows: In section 2, related works are briefly discussed. The proposed approach is presented in section 3. Results are given in section 4. Analysis and discussion are given in section 5. Finally, conclusions are presented in section 6.

2. RELATED WORKS

ZM & PZM descriptors can be utilized for obtaining a quantitative shape feature description of the hand for recognizing static gestures. In general, ZMs are better in describing shapes than PZMs, whereas PZMs are less affected by noise than ZMs.

The 2D complex ZMs of order n with repetition m of an image, $I(\rho, \theta)$, in the polar coordinates are defined inside the unit circle as

$$A_{nm} = \frac{n+1}{\pi}\sum_{\rho}\sum_{\theta}[V_{nm}(\rho, \theta)]^{*} \cdot I(\rho, \theta), \ \rho \leq 1, \tag{1}$$

where $n = 0, 1, 2, \ldots, \infty$, $|m| \leq n$, and $n - |m|$ is even. The symbol * denotes the complex conjugate. The form of orthogonal Zernike basis polynomials, $V_{nm}(\rho, \theta)$ is defined as

$$V_{nm}(\rho, \theta) = R_{nm}(\rho) \cdot \exp(-jm\theta), \tag{2}$$

where the radial polynomials $R_{nm}(\rho)$ is defined as

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s!((n+|m|)/2-s)!((n-|m|)/2-s)!}\rho^{n-2s} \tag{3}$$

and $\exp(-jm\theta)$ is a phase factor.

Pseudo-Zernike basis polynomials are another set of orthogonal basis polynomials with similar properties to Zernike basis polynomials. The pseudo-Zernike basis polynomials differ from the Zernike ones in that the radial polynomials are defined as

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|} (-1)^s \frac{(2n+1-s)!}{s!(n-|m|-s)!(n+|m|+1-s)!}\rho^{n-s} \tag{4}$$

where $n = 0, 1, 2, \ldots, \infty$, and $|m| \leq n$. Similarly, the 2D complex PZMs can be defined in the same way as Eq. (1).

ZMs & PZMs are defined inside the unit circle. They are descriptors for globally describing hand shapes but do not have reliable local support. Besides, conventionally, the original is moved to the centroid of the hand shape to achieve translation invariance.

However, the centroid of the hand shape would be unstable while erroneous noises occur. The computed ZMs & PZMs would not be robust.

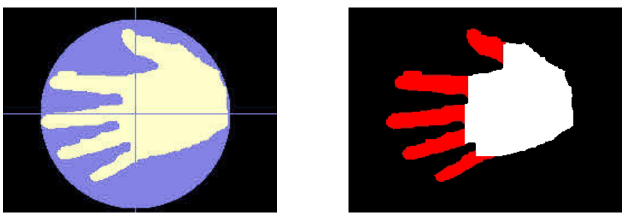
Kim *et al*. [9] imposed weightings on shapes prior to computing the ZMs such that the inner and outer shape properties could be reflected in the ZMs. This imposition of weightings is only suitable for symmetrical shapes. However, hand shapes are not symmetrical and imposing shape weightings is not much help.

3. THE PROPOSED APPROACH

The primary objectives of this study are to present a new approach for static gesture recognition. The steps of the proposed approach are as follows.

Algorithm NewSGR
Input: a static gesture image.
Output: the recognized static gesture image.
Step 1: Segment the input static gesture image into the binary hand silhouette.
Step 2: Accommodate the binary hand silhouette with a minimum bounding circle (MBC).
Step 3: Decompose the binary hand silhouette into the finger part and the palm part by morphological operations according to the radius of the MBC.
Step 4: Compute the ZMs & PZMs of the finger part and the palm part with different importance, respectively, based on the center of the MBC.
Step 5: Calculate distances between the input feature vector and the stored feature vectors.
Step 6: Output the stored static gesture with the minimum distance.

In the first step, an input static gesture image is segmented into a binary hand silhouette via the modified color segmentation approach [4]. In the second step, the binary hand silhouette is accommodated with an MBC. In the third step, the binary hand silhouette is decomposed into the finger part and the palm part by morphological operations according to the radius of the MBC. Fig. 2 (a) shows an example of an MBC that accommodates the inside of the binary hand silhouette; Fig. 2 (b) shows the finger part and palm part of the binary hand silhouette.



(a) (b)

Fig. 2. (a) MBC that accommodates the binary hand silhouette inside and (b) finger part and palm part of the binary hand silhouette.

In the fourth step, ZMs & PZMs of the finger part and the palm part with different importance, respectively, are computed based on the center of the MBC. Let F^I be the feature vector of the input static gesture *I*. The ZMs & PZMs of the input static gesture *I* are defined by

$$F^I = (FEA^I_{finger}[i],\ FEA^I_{palm}[i]), \tag{5}$$

where $FEA^I_{finger}[i]$ and $FEA^I_{palm}[i]$ are the *i*-th ZMs & PZMs of the finger part and the palm part, respectively, of an input static gesture *I*. The importance-driven ZMs & PZMs of the input static gesture *I* are then defined by

$$F^I = (W_{finger} FEA^I_{finger}[i],\ W_{palm} FEA^I_{palm}[i]), \tag{6}$$

where W_{finger} and W_{palm} are weights of the finger part and the palm part, respectively.

Let F^S be the feature vector of a stored image *S*. The ZMs & PZMs of the stored static gesture *S* are defined by

$$F^S = (FEA^S_{finger}[i],\ FEA^S_{palm}[i]), \tag{7}$$

where $FEA^S_{finger}[i]$ and $FEA^S_{palm}[i]$ are the *i*-th ZMs & PZMs of the finger part and the palm part, respectively, of a stored static gesture *S*. The importance-driven ZMs & PZMs of the input static gesture *S* are then defined by

$$F^S = (W_{finger} FEA^S_{finger}[i],\ W_{palm} FEA^S_{palm}[i]). \tag{8}$$

In the fifth step, distances between the input feature vector and the stored feature vectors are calculated. The distance function between the input feature vector and the stored feature vector is defined by

$$Dist(F^I, F^S) = W_{finger} \cdot Dist_{finger}(F^I, F^S) + W_{palm} \cdot Dist_{palm}(F^I, F^S), \tag{9}$$

where $Dist_{finger}(F^I, F^S)$ and $Dist_{palm}(F^I, F^S)$ are defined as

$$Dist_{finger}(F^I,\ F^S) = \sum_{i=0}^{N_{finger}-1} \| FEA^I_{finger}[i] - FEA^S_{finger}[i] \| \tag{10}$$

and

$$Dist_{palm}(F^I,\ F^S) = \sum_{i=0}^{N_{palm}-1} \| FEA^I_{palm}[i] - FEA^S_{palm}[i] \|, \tag{11}$$

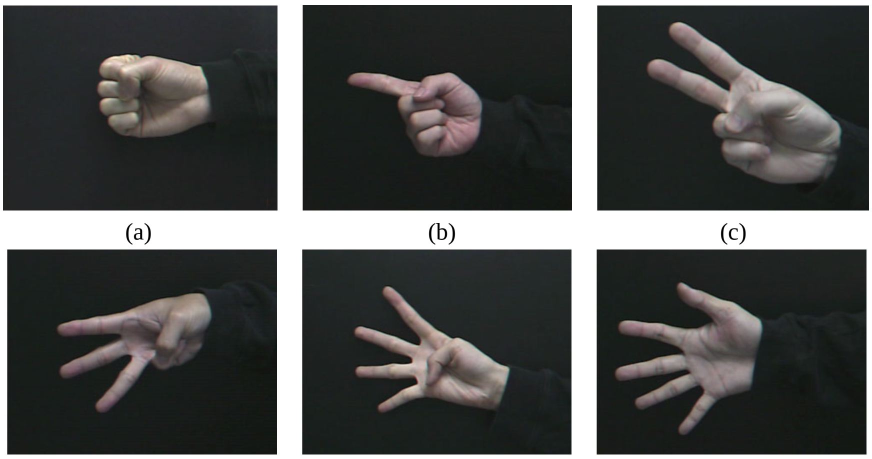
respectively, where N_{finger} and N_{palm} are the numbers of ZMs & PZMs of the finger part and the palm part, respectively.

In the final step, 1-nearest neighbor techniques are used for static gesture identification. The stored static gesture with the minimum distance is output.

4. RESULTS

To evaluate the effectiveness of the proposed approach to static gesture recognition, several experiments were tested.

For the purpose of recognizing a static gesture feature represented by feature parameter F^I, a database, including a plurality of feature parameters F^S corresponding to reference static gestures, was provided. In general, a static gesture can be used to represent a certain digit. For example, the index finger represented digit "1," the index finger plus middle finger represented digit "2," the five fingers of a hand represented digit "5," and so on. The experimental images were 320 × 240 pixel color images of 6 static gestures, namely, Zero, One, Two, Three, Four, and Five, as shown in Fig. 3. The database consists of 6 different static gestures formed by 10 subjects. Each subjects' static gesture was captured 10 times, each time with different translation, scale and rotation. The total number of static gesture images in the database was 600. The images were divided into two sets: 1) a stored set containing 300 images and 2) a testing set containing 300 images. The experimental platform was a PC with an Intel® Pentium® 4 CPU 2.80GHz CPU running Microsoft® Windows XP Professional OS.



(a) (b) (c)

(d) (e) (f)

Fig. 3. Six static gestures, namely, zero, one, two, three, four, and five, respectively.

The effects on the recognition rate of the proposed approach with different values of pair (W_{finger}, W_{palm}) were done in the first experiment. The order *n* was fixed to be 7. Table 1 displays the results of the proposed approach with different values of weights. In this table, different values of pair (W_{finger}, W_{palm}) are displayed in the first column; the second column shows the recognition rate of the proposed approach. From Table 1, the best recognition rate when pair (W_{finger}, W_{palm}) was set at (0.7, 0.3) was 97.3%; the worst recognition rate when pair (W_{finger}, W_{palm}) was set at (0.9, 0.1) or (0.5, 0.5) was 95.0%.

The effects on the recognition rate with different values of order *n* were done in the second experiment. It was compared the previous method with the proposed approach.

Table 1. Results of the proposed approach with different values of weights.

(W_{finger}, W_{palm})	Recognition Rate (%)
(0.9, 0.1)	95.0
(0.8, 0.2)	97.0
(0.7, 0.3)	97.3
(0.6, 0.4)	96.3
(0.5, 0.5)	95.0

Table 2. Comparison of the previous method and the proposed approach with different values of orders.

n	Recognition Rate (%)	
	Previous Method	Proposed Approach
6	95.0	97.0
7	95.0	97.3
8	95.0	97.3
9	94.7	97.0
10	94.7	97.0

The pair (W_{finger}, W_{palm}) was fixed to be (0.7, 0.3). Table 2 displays the results of the previous method and the proposed approach with different values of orders. In this table, different values of order n are displayed in the first column; the second column shows the recognition rate of the previous method; the last column shows the recognition rate of the proposed approach. From Table 2, the best recognition rate of the proposed approach when order n was set at 7 or 8 was 97.3%; the worst recognition rate of the proposed approach when order n was set at 6, 9 or 10 was 97.0%. For each order n, the recognition rate of the proposed approach is better than that of the previous method. Hence, the proposed approach does improve the recognition rate of static gestures.

5. ANALYSIS AND DISCUSSION

This paper has proposed a new algorithm for static gesture recognition based on ZMs & PZMs.

Based on the center of the MBC, the projections of hand silhouettes on Zernike and pseudo-Zernike bases could be computed accurately. Besides, the radius of the MBC can be used to accurately decompose the binary hand silhouette into the finger part and the palm part by morphological operations. Hence, the proposed approach can extract reliable hand features.

It is not easy to determine the appropriate parameters for the proposed. From Table 1, pair (W_{finger}, W_{palm}) was set at (0.7, 0.3) was can get the best recognition rate, heuristically. Since the finger part of the hand-shape plays deformable role for determining the hand-shape, more importance to the finger part of static gestures is reasonable. From Table 2, the order set at 7 for the proposed approach can achieve a trade-off between computation complexity and accuracy. Higher orders would not contribute to the recognition rate but to the computation cost.

The results show that the proposed approach performs well in recognizing the 6 static gestures. However, there are some instances of failure for the proposed approach. After analyzing the negative results of the proposed approach, the main reasons for the negatives are as follows. First, two adjacent fingers of a static gesture are too close and the segmented hand silhouette would be undesirable. Hence, this leads to the negative results. For example, Fig. 4 shows (a) input static gesture and (b) the undesirable silhouette of the input. Second, some static gestures are similar. The distance between the features of the similar silhouettes would be the smallest sometimes. For example, some static gestures Five recognize some static gestures Four.

(a) (b)

Fig. 4. (a) Input static gesture and (b) the undesirable silhouette of the input.

6. CONCLUSIONS

A new approach for static gesture recognition based on ZMs & PZMs has been presented. The main contributions of this paper are as follows: a) using the center of the MBC for translation invariance, more robust static gesture features can be obtained; b) the proposed approach extracts weighted ZMs & PZMs of static gestures, thereby describing static gestures globally with local support and providing a reliable description of static gesture features; c) the proposed approach performs better than previous methods based on conventional ZMs & PZMs.

This approach takes a step toward extracting reliable features for static gesture recognition. Experiments similar to those reported here would be done using a wider variety of static gestures in future research.

ACKNOWLEDGMENTS

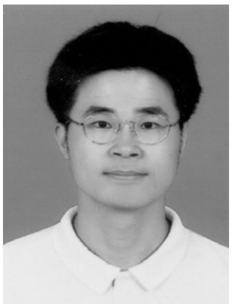
This work is a partial result of project No. 3XS1B11 conducted by ITRI under the sponsorship of the Ministry of Economic Affairs, Taiwan, R.O.C.

REFERENCES

1. D. S. Banarse and A. W. G. Duller, “Deformation invariant pattern classification for

recognizing hand gestures," in *Proceedings of the IEEE International Conference on Neural Networks*, Vol. 3, 1996, pp. 1812-1817.
2. C. C. Chang, Y. S. Huang, and J. J. Chen, "Importance-driven recognition of static gestures," in *Proceedings of the International Workshop on Advanced Image Technology* (*IWAIT*), 2002, pp. 163-166.
3. J. W. Deng and H. T. Tsui, "A PCA/MDA scheme for hand posture recognition," in *Proceedings of the 16th International Conference on Pattern Recognition*, Vol. 1, 2002, pp. 283-286.
4. P. Felzenszwalb and D. Huttenlocher, "Image segmentation using local variation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1998, pp. 98-104.
5. E. Gose, R. Johnsonbaugh, and S. Jost, *Pattern Recognition and Image Analysis*, Prentice Hall, 1996.
6. T. S. Huang and V. I. Pavlovic, "Hand gesture modeling, analysis and synthesis," in *Proceedings of the International Workshop on Automatic Face and Gesture Recognition*, 1995, pp. 73-79.
7. E. Hunter, J. Schlenzig, and R. Jain, "Posture estimation in reduced-model gesture input systems," in *Proceedings of the International Workshop on Automatic Face and Gesture Recognition*, 1995, pp. 290-295.
8. A. Khotanzad and Y. H. Hong, "Rotational invariant image recognition using features selected via a systematic method," *Pattern Recognition*, Vol. 23, 1990, pp. 1089-1101.
9. H. K. Kim, J. D. Kim, D. G. Sim, and D. I. Oh, "A modified Zernike moment shape descriptor invariant to translation, rotation and scale for similarity-based image retrieval," in *Proceedings of the IEEE International Conference on Multimedia and Expo*, Vol. 1, 2000, pp. 307-310.
10. H. J. Lee and J. H. Chung, "Hand gesture recognition using orientation histogram," in *Proceedings of the IEEE Region 10 Conference* (*TENCON*), Vol. 2, 1999, pp. 1355-1358.
11. C. Nolker and H. Ritter, "Visual recognition of continuous hand postures," *IEEE Transactions on Neural Networks*, Vol. 13, 2002, pp. 983-994.
12. V. I. Pavlovic, R. Sharma, and T. S. Huang, "Visual interpretation of hand gestures for human-computer interaction: a review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, 1998, pp. 496-513.
13. F. K. H. Quek and M. Zhao, "Inductive learning in hand pose recognition," in *Proceedings of the 2nd International Conference on Automatic Face and Gesture Recognition*, 1996, pp. 78-83.
14. J. Schlenzig, E. Hunter, and R. Jain, "Vision based hand gesture interpretation using recursive estimation," in *Proceedings of the 28th Asilomar Conference on Signals, Systems and Computers*, Vol. 2, 1994, pp. 1267-1271.
15. M. C. Su, W. F. Jean, and H. T. Chang, "A static hand gesture recognition system using a composite neural network," in *Proceedings of the 5th IEEE International Conference on Fuzzy Systems*, Vol. 2, 1996, pp. 786-792.
16. C. H. Teh and R. T. Chin, "On image analysis by the methods of moments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, 1988, pp. 496-513.

17. J. Triesch and C. von der Malsburg, "A system for person-independent hand posture recognition against complex backgrounds," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, 2001, pp. 1449-1453.
18. Y. Wu and T. S. Huang, "View-independent recognition of hand postures," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Vol. II, 2000, pp. 88-94.
19. S. G. Wysoski, M. V. Lamar, S. Kuroyanagi, and A. Iwata, "A rotation invariant approach on static-gesture recognition using boundary histograms and neural networks," in *Proceedings of the 9th International Conference on Neural Information Processing* (*ICONIP*), Vol. 4, 2002, pp. 2137-2141.
20. C. T. Zhan and R. Z. Roskies, "Fourier descriptor for plane closed curves," *IEEE Transactions on Computer*, Vol. 21, 1972, pp. 269-281.
21. H. Zhou, D. J. Lin, and T. S. Huang, "Static hand gesture recognition based on local orientation histogram feature distribution model," in *Proceedings of the IEEE Workshop on Real-Time Vision for Human Computer Interactions in Conjunction with CVPR04*, 2004, pp. 161-164.



Chin-Chen Chang (張勤振) received his Ph.D. degree in Computer Science from National Chiao Tung University in Hsinchu, Taiwan, R.O.C., in 1998. At present, he is an Assistant Professor in the Department of Computer Science and Information Engineering, National United University in Miaoli, Taiwan, R.O.C. His research interests include computer graphics, human-computer interface, and multimedia systems.

Jiann-Jone Chen (陳建中) was born in Taichung, Taiwan in 1966. He received the B.S. and M.S. degrees from National Cheng Kung University, Tainan, in 1989 and 1991, respectively, and the Ph.D. from National Chiao Tung University, Hsinchu, Taiwan, in 1997, all in Electrical Engineering. From 1999 to 2002, he was with CCL/ITRI, where he was engaged in digital image processing and multimedia indexing. He joined the Electrical Engineering Department of National Taiwan University of Science and Technology, Taipei, Taiwan, in 2002. His current research interests include digital video compression, image/video processing and multimedia indexing and computer graphics.

Wen-Kai Tai (戴文凱) was born on January 29, 1965, in Hualien, Taiwan, R.O.C. He received his M.S. and Ph.D. degrees in Computer Science from National Chiao Tung University in 1989 and 1995, respectively. He is currently an Associate Professor of Computer Science and Information Engineering Department at National Dong Hwa University, Taiwan, R.O.C. His research interests include texture synthesis, visibility, and gaming.

Chin-Chuan Han (韓欽銓) received the B.S. degree in Computer Engineering from National Chiao Tung University in 1989, and an M.S. and a Ph.D. degrees in Computer Science and Electronic Engineering from National Central University in 1991 and 1994, respectively. From 1995 to 1998, he was a Postdoctoral Fellow in the Institute of Information Science, Academia Sinica, Taipei, Taiwan. He was an Assistant Research Fellow in the Applied Research Lab., Telecommunication Laboratories, Chunghwa Telecom Co. in 1999. From 2000 to 2004 he was an Associate Professor in the Department of Computer Science and Information Engineering, Chung Hua University. He is currently an Associate Professor in the Department of Computer Science and Information Engineering, National United University. He is a member of IEEE, SPIE, TAAI, and IPPR in Taiwan. His research interests are in the areas of face recognition, biometrics authentication, image analysis, computer vision, and pattern recognition.