行政院國家科學委員會專題研究計畫 成果報告

移動平臺之前瞻性視訊監控技術
研究成果報告(精簡版)

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 97-2221-E-216-040-
執 行 期 間 ：97 年 08 月 01 日至 98 年 09 月 30 日
執 行 單 位 ：中華大學資訊工程研究所

計 畫 主 持 人 ：黃雅軒

報 告 附 件 ：出席國際會議研究心得報告及發表論文

處 理 方 式 ：本計畫可公開查詢

中 華 民 國 98 年 10 月 29 日

附件一

行政院國家科學委員會補助專題研究計畫 ■成果報告 □期中進度報告

移動平臺之前瞻性視訊監控技術

計畫類別：■ 個別型計畫　　□ 整合型計畫
計畫編號：NSC　97-2221-E-216-040
執行期間：　97　年　8　月　1　日至　98　年　9　月　30　日

計畫主持人：黃雅軒
共同主持人：
計畫參與人員：　莊順旭、許廷嘉、王勻駿

成果報告類型(依經費核定清單規定繳交)：■精簡報告　□完整報告

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
■出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢
□涉及專利或其他智慧財產權，□一年■二年後可公開查詢

執行單位：中華大學

中　華　民　國　九十八　年　十　月　二十八　日

附件二

可供推廣之研發成果資料表

□ 可申請專利　■ 可技術移轉　日期：98年10月28日

國科會補助計畫	計畫名稱：移動平臺之前瞻性視訊監控技術 計畫主持人：黃雅軒 計畫編號：NSC 97-2221-E-216-040 學門領域：資訊
技術/創作名稱	強健性人臉辨識
發明人/創作人	黃雅軒
技術說明	● 中文：本技術實現二種具高鑑別度的人臉辨識方法(CMSM 和 GDA)，並將他們有效的組合，以得到一套具強鍵性的人臉辨識系統。CMSM (Constrained Mutual Subspace Method，限制性子空間)使用多張影像所形成的子空間作為辨識的依據，可表示使用者人臉特有的變化形態。GDA (Generalized Discriminant Analysis)使用核函數之非線性區別分析，將資料映射至高維度空間使其盡量線性可分割。這二種方法使用不同種類的特徵和比對機制，因此他們的比對結果具有高度的互補性。針對通用的 Banca 人臉資料庫，本研究的辨識結果在容忍 10%的錯誤接受率(False Accept Rate)條件下，正確的辨識率(Recognition Rate)可高達 98%。
	英文：*This paper presents a robust face recognition method which two highly discriminating algorithms (CMSM and GDA) to recognize human faces. CMSM (Constraint Mutual Subspace Method) constructs a class subspace for each person and makes the relation between class subspaces by projecting them onto a generalized difference subspace so that the canonical angles between subspaces are enlarged to approach to the orthogonal relation. GDA (Generalized Discriminant Analysis) adopts kernel function operator to make it easy to extend and generalize the classical Linear Discriminant Analysis to a non linear one. Both CMSM and GDA are effective to recognize human faces, however, CMSM constructs a subspace from several face images and GDA needs only one face image to perform recognition. Obviously, these two methods inherently have different properties and abilities of recognition so that we combine them together. Experimental results show that the proposed method can achieve good recognition accuracy.*

可利用之產業 及 可開發之產品	可利用之產業：安全監控、機器人、系統整合、互動遊戲 可開發之產品：智慧型門禁系統、智慧機器人、互動遊戲
技術特點	● 對光線變化具有容忍能力 ● 即時處理 ● 正確率高 ● 應用範圍廣
推廣及運用的價值	可判斷使用者的身份，增加生活上使用不同產品的便利性和安全性。另外，針對互動遊戲亦可增加其趣味性。

※ 1. 每項研發成果請填寫一式二份，一份隨成果報告送繳本會，一份送 貴單位研發成果推廣單位（如技術移轉中心）。

※ **2. 本項研發成果若尚未申請專利，請勿揭露可申請專利之主要內容。**

※ 3. 本表若不敷使用，請自行影印使用。

計畫成果：

本計畫的研發成果已發表了四篇論文：

1. Yea-Shuan Huang, Wei-Cheng Liu and Fang-Hsuan Cheng, “Face Recognition by Combining Complementary Matchings of Single Image and Sequential Images”, MVP 2009 IAPR Conference on Machine Vision Applications, pp.253~256, 2009.
2. Ting-Chia Hsu, Yea-Shuan Huang, Shun-Hsu Chuang, Yun-Jiun Wang and Shian Wan, “An Improved ASM-Based Facial Feature Locating Method”, CVGIP. 2009.
3. Yea-Shuan Huang, Wei-Cheng Liu and Shian Wan, “Improvement of The Constrained Mutual Subspace Method for Face Recognition”, 2009.
4. Yea-Shuan Huang and Wei-Cheng Liu, “Face Recognition Based on Complementary Matching of Single Image and Sequential Images”, IIHSMP, 2009.

Face Recognition Based on Complementary Matching of Single Image and Sequential Images

Yea-Shuan Huang and Wei-Cheng Liu
Computer Science & Information Engineering, Chung Hua University
No.707, Sec. WuFu Rd., Hsinchu, Taiwan, 300, R.O.C.

Abstract

This paper presents a robust face recognition method which two highly discriminating algorithms (CMSM and GDA) to recognize human faces. CMSM (Constraint Mutual Subspace Method) constructs a class subspace for each person and makes the relation between class subspaces by projecting them onto a generalized difference subspace so that the canonical angles between subspaces are enlarged to approach to the orthogonal relation. GDA (Generalized Discriminant Analysis) adopts kernel function operator to make it easy to extend and generalize the classical Linear Discriminant Analysis to a non linear one. Both CMSM and GDA are effective to recognize human faces, however, CMSM constructs a subspace from several face images and GDA needs only one face image to perform recognition. Obviously, these two methods inherently have different properties and abilities of recognition so that we combine them together. Experimental results show that the proposed method can achieve good recognition accuracy.

1. Introduction

Biometric identification technology is a very popular research field in the recent years. Various methods have been proposed that use different kinds of biometric data. Among them, face recognition consistently obtains a great expectation since it is contact-free and is user friendly. Therefore, a lot of research efforts have been devoted to this field, and many face recognition approaches based on a variety of machine learning theorem have been developed already. For example, subspace methods such as PCA [1] and LDA [2] are commonly used which project high dimensional features to low dimensional features and not only faster but also better recognition can be achieved. In general, LDA has better recognition ability than PCA which is based on an eigenvalue resolution and gives an exact solution of the maximum of the inertia. But even LDA fails for a nonlinear problem. A Generalized Discriminant Analysis (GDA) is developed to overcome this difficulty by mapping the input space into a high dimensional feature space with linear properties so that it can solve the problem in a LDA classical way.

Basically, the feature derived from a single image denotes the location of this image in a high dimensional feature space. In the feature space, the locations corresponding to two similar images will be in general close to each other, and the locations of two very distinct images then will be quite separated apart. Therefore, recognition based on a single image mainly measures the distance (or similarity) of the features between the input pattern and the reference patterns. However, the feature derived from a set of sequential images of the same person can present the unique variation model of this person. Therefore, recognition based on sequential images indeed compares the specific variation pattern of the unknown input subject and that of each individual class. The two kinds of recognition seem to be complementary in nature. With this understanding, it will be very useful if both kinds of methods are combined together. In this paper, GDA (Generalized Discriminant Analysis) and CMSM (Constraint Mutual Subspace Method) are used together to recognize faces not only because they both have high recognition abilities, but also they probably are complementary to each other since GDA takes a single-image matching strategy and CMSM takes a sequential-image matching strategy.

This paper is organized as follows. Section 2 describes two recognition models; the first is GDA and the second is CMSM. A linear mechanism is also proposed to integrate their recognition results. Section 3 presents the experiment results on the famous Banca face database, and the final conclusion is drawn in Section 4.

2. Face identification method

In this section, we describe our face recognition framework which integrates a single-image matching module and a sequence-image matching module. The single-image matching module uses a GDA algorithm to reduce feature dimension first, and then adopts a nearest distance classification to recognize the input pattern; the sequence-image matching module uses a CMSM metric which projects each individual subspace including the input and the reference subspaces onto a common difference subspace and their canonical angles are used to recognize the input patterns. For making the final decision, a linear combination scheme is used to integrate the two matching scores. This section consists of three subsections. Subsection 2.1 states the matching of a single image which adopts the Euclidean distance in a GDA-transformed reduced feature space. Subsection 2.2 describes the CMSM algorithm containing the construction of Constrain mutual subspace, computation of canonical angle and matching. Finally, Subsection 2.3 describes the weighted-sum combination scheme.

2.1. Generalized discriminant analysis

Linear Discriminant Analysis (LDA) is a traditional statistical method which has been proven successful on classification problems, however it will fail to deal with nonlinear problems. Therefore, Generalized Discriminant Analysis (GDA) is proposed to overcome this situation by mapping the input space into a convenient feature space in which variables are nonlinearly related to the input space. In the following of this section, the notations and the formulation of the GDA using dot product and matrix form are explained.

Let L be the total number of classes and N_i be the number of training samples belonging to class i, x_j^i be the jth sample of class i, $\phi(x_j^i)$ be a nonlinear mapping of x_j^i into a high-dimensional Hilbert feature space, $X_i^T = [\phi(x_1^i), \dots, \phi(x_{N_i}^i)]$ and $X^T = [X_1^T, \dots, X_L^T]$. Suppose there are in total N training samples, i.e. $\{x_1, x_2, \cdots, x_N\}$, there is a kernel matrix K of which each components is the inner product value of the high-dimensional mapping features of two samples. That is

$$K = \begin{bmatrix} \langle\phi(x_1)\cdot\phi(x_1)\rangle & \cdots & \langle\phi(x_1)\cdot\phi(x_N)\rangle \\ \vdots & \ddots & \vdots \\ \langle\phi(x_N)\cdot\phi(x_1)\rangle & \cdots & \langle\phi(x_N)\cdot\phi(x_N)\rangle \end{bmatrix} = \begin{bmatrix} \kappa(x_1,x_1) & \cdots & \kappa(x_1,x_N) \\ \vdots & \ddots & \vdots \\ \kappa(x_N,x_1) & \cdots & \kappa(x_N,x_N) \end{bmatrix}$$

In general, RBF (Radial Basis Function) kernel can be chosen to serve as κ as

$$\kappa(x_i, x_j) = exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right), \sigma \in R - \{0\}$$

By changing the value of σ, the most appropriate feature space can be constructed. Let

$$W_i = \begin{bmatrix} \frac{1}{N_i^2} & \cdots & \frac{1}{N_i^2} \\ \vdots & \ddots & \vdots \\ \frac{1}{N_i^2} & \cdots & \frac{1}{N_i^2} \end{bmatrix}, \qquad W = \begin{bmatrix} N_1 W_1 & & 0 \\ & \ddots & \\ 0 & & N_L W_L \end{bmatrix},$$

$$m_i = \frac{1}{N_i}\sum_{j=1}^{N_i} \phi\left(x_j^{(i)}\right) = \left[\phi\left(x_1^{(i)}\right), \dots, \phi\left(x_{N_i}^{(i)}\right)\right] \begin{bmatrix} \frac{1}{N_i} \\ \vdots \\ \frac{1}{N_i} \end{bmatrix} = X_i^T \begin{bmatrix} \frac{1}{N_i} \\ \vdots \\ \frac{1}{N_i} \end{bmatrix}$$

$$m_i m_i^T = X_i^T \begin{bmatrix} \frac{1}{N_i} \\ \vdots \\ \frac{1}{N_i} \end{bmatrix} \left[\frac{1}{N_i}, \dots, \frac{1}{N_i}\right] X_i = X_i^T W_i X_i.$$

Suppose the high-dimensional features of all samples are already centered at the original point (i.e. their mean value m_0 is 0), then the between-class scatter matrix S_b^{GDA} and the within-class scatter matrix S_b^{GDA} are defined as

$$\begin{aligned} S_b^{GDA} &= \sum_{i=1}^{L} \frac{N_i}{N}(m_i - m_0)(m_i - m_0)^T \\ &= \sum_{i=1}^{L} \frac{N_i}{N} m_i m_i^T \\ &= \frac{1}{N}\sum_{i=1}^{L} N_i X_i^T W_i X_i \\ &= \frac{1}{N}[X_1^T, \dots, X_L^T] \begin{bmatrix} N_1 W_1 & & 0 \\ & \ddots & \\ 0 & & N_L W_L \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ X_L \end{bmatrix} \\ &= \frac{1}{N} X^T W X \end{aligned}$$

and

$$\begin{aligned} S_w^{GDA} &= \frac{1}{N}\sum_{i=1}^{L}\sum_{j=1}^{N_i} \phi\left(x_j^{(i)}\right)\phi\left(x_j^{(i)}\right)^T \\ &= \frac{1}{N}\sum_{i=1}^{L}\left[\phi\left(x_1^{(i)}\right), \dots, \phi\left(x_{N_i}^{(i)}\right)\right] \begin{bmatrix} \phi\left(x_1^{(i)}\right) \\ \vdots \\ \phi\left(x_{N_i}^{(i)}\right) \end{bmatrix} \\ &= \frac{1}{N}\sum_{i=1}^{L} X_i^T X_i = \frac{1}{N}[X_1^T, \dots, X_L^T] \begin{bmatrix} X_1 \\ \vdots \\ X_L \end{bmatrix} \\ &= \frac{1}{N} X^T X \end{aligned}$$

The object of GDA is to find the transformed vector v which gains the largest ratio between S_b^{GDA} and S_w^{GDA} in the transformed space，that is

$$v = \arg\max \frac{v^T S_b^{GDA} v}{v^T S_w^{GDA} v}$$

Now, this becomes the Eigen problem of finding $S_b^{GDA} v = \lambda S_w^{GDA} v$。From Linear Algebra, the transform vector v can be derived as a linear combination of the corresponding high-dimensional mapping vectors of the collected samples, that is $v = X^T\alpha$. Therefore, taking $S_b^{GDA} = \frac{1}{N}X^T W X$ and $S_w^{GDA} = \frac{1}{N}X^T X$, the above equation becomes

$$\begin{aligned} & S_b^{GDA} v = \lambda S_w^{GDA} v \\ \Rightarrow\ & \left(\tfrac{1}{N}X^T W X\right) v = \lambda\left(\tfrac{1}{N}X^T X\right) v \\ \Rightarrow\ & X^T W X X^T \alpha = \lambda X^T X X^T \alpha \qquad \ldots(v = X^T\alpha) \\ \Rightarrow\ & X X^T W X X^T \alpha = \lambda X X^T X X^T \alpha \\ \Rightarrow\ & (KWK)\alpha = \lambda(KK)\alpha \qquad \ldots(K = XX^T) \end{aligned}$$

Although the Feature Mapping ϕ is unknown so that X cannot be computed directly, K and W in fact are computable. Therefore, by using the generalized Eigen problem solving method, the eigenvectors α corresponding to the large eigenvalues λ can be derived first and then the transform vectors v can also be derived. The kernel operator K allows the construction of nonlinear separating function in the input space that is equivalent to linear separating function in the feature space F. Through the kernel method, GDA in general has much better discrimination ability than LDA.

The transformed feature y now becomes $y = v^T x$ where x is a sample feature vector. For recognition, a nearest distance classification metric is applied. Let $I_1, \cdots, I_m$ denote the feature vectors of m input samples, $I_1^{GDA}, \ldots, I_m^{GDA}$ are their GDA transformed vectors, and $R_{k,q}^{GDA}$ $(q = 1, \ldots, n)$ be the GDA transformed feature vector of the qth sample of the kth enrolled person. Then, the distance of the m input samples and the n reference data of the kth person becomes

$$Dist(k) = \min_{p=1,\cdots,m} \min_{q=1,\cdots,n} d\left(I_p^{GDA}, R_{k,q}^{GDA}\right)$$

2.2. Constrain mutual subspace method

2.2.1. Concept of canonical angle

In linear algebra, the similarity between two subspaces is calculated by the angle between them. Suppose $\{R_1, \ldots, R_r\}$ is a set of r reference patterns, $\{I_1, \ldots, I_s\}$ is a set of s input patterns, and each pattern is represented by an f-dimensional feature vector. With PCA, an r_{no}-dimensional reference subspace Ω can be constructed from $\{R_1, \ldots, R_r\}$, and an s_{no}-dimensional input subspace Λ can be constructed from $\{I_1, \ldots, I_s\}$ respectively. Therefore, Ω is an $r_{no} \times f$ matrix and Λ is an $s_{no} \times f$ matrix. In general, the relations of r, s, r_{no} and s_{no} are chosen to be $r_{no} \leqq r$, $s_{no} \leqq s$ and $r_{no} \leqq s_{no}$. We can further obtain r_{no} canonical angles $\{\theta_1, \ldots, \theta_{r_{no}}\}$ between subspace Ω and subspace Λ by the following equations:

$$XC = \lambda C$$

$$X = (x_{ij}),\ x_{ij} = \sum_{k=1}^{r_{no}} (\psi_i \cdot \phi_k)(\phi_k \cdot \psi_j)$$

where ψ_i and ϕ_i denote respectively the i-th f-dimensional orthonormal basis vector of subspace Ω and Λ, λ is an eigenvalue of X and C is the eigenvectors of X, and X is an $r_{no} \times r_{no}$ matrix. The value $cos^2\, \theta_i$ of the i-th smallest canonical angle equals to the i-th largest eigenvalue of Λ. The largest eigenvalue (i.e. $cos^2\, \theta_1$) is taken to denote the similarity between subspace Ω and Λ.

2.2.2. Generation of constrained subspace

In CMSM, it is essentially important to generate a proper constrained subspace C which contains the effective matching components but eliminating the unnecessary ones. By projecting the input subspace and reference subspaces to a constrained subspace, it could extract discriminating features for recognizing pattern classes.

Suppose there are in total N_p reference subspaces. To generate a constrained subspace, we compute the projection matrix Ω_k of the k-th reference subspace using

$$P_k = \sum_{j=1}^{r_{no}} \psi_j^k \left(\psi_j^k\right)^T$$

where r_{no} is the number of eigenvectors of a reference subspace, ψ_j^k is the j-th orthonormal basis vector of the k-th reference subspace, and each P_k is a $f \times f$ matrix. Then, we calculate the eigenvectors of the summation matrix $S = \left(P_1 + P_2 + \cdots + P_{N_p}\right)$, that is $SA = \lambda A$, where λ and A denote the eigenvalues and the eigenvectors of S respectively. Finally, the t eigenvectors $[A_1, \ldots, A_t]$ corresponding to the t smallest eigenvalues are selected to construct the constrained subspace CS (that is $CS = [A_1, \ldots, A_t]_{t \times f}$). For a more detailed description of CMSM, please see [4].

2.2.3. Matching on constrained subspace

Suppose there are in total L recognition classes. Π denotes the input subspace derived from the input sequence samples, and T_k $(1 \leqq k \leqq L)$ denotes the subspace derived from the training sequence samples of class k. Five steps need to be performed for pattern matching as follows:

1. Project each T_k onto CS and generate an $r_{no} \times t$ projection matrix P_k;
2. Normalize each P_k, and with a Gram-Schmidt algorithm derive a reference subspace Ω_k with basis $\{\psi_1^k, \cdots, \psi_{t\_no}^k\}$;
3. Project Π onto CS and generate an $s_{no} \times t$ projection matrix Q;
4. Normalize Q, and with a Gram-Schmidt algorithm derive the input subspace Λ with basis $\{\phi_1, \cdots, \phi_{t\_no}\}$;
5. Compute the similarity $Sim(k)$ between Λ and Ω_i by using the canonical angle computation as

$$Sim(k) = \sum_i^{s_{no}} \sum_j^{r_{no}} (\psi_i^k, \phi_j)^2$$

2.3. Combination scheme

Obviously, $Dist(k)$ is a distance measurement and $Sim(k)$ is a similarity measurement, they have totally different interpretation, and both small $Dist(k)$ and large $Sim(k)$ denote that the input patterns and the reference data of person k are similar to each other. In order to combine the matching scores of GDA and CMSM, the integrated value of similarity is calculated as

$$similarity(k) = \omega_1 \times Sim(k) + \omega_2 \times \left(\alpha - \frac{Dist(k)}{\sigma}\right)$$

where ω_1 and ω_2 are the combining weights of the two matching scores, and α and σ are two normalized parameter. All the parameters are decided by experiments.

3. Experimental results

We used the famous Banca face database to evaluate the performance of the proposed recognition method. The Banca database contains 52 individuals and each individual has 12 image sequences that were taken in different time, at different locations and by difference cameras. Each image sequence consists of 10 face images with various facial poses and facial expressions. To simplify the problem, only 4 image sequences of each individual taken in different time at the same locations with the same camera are used in this experiment. Among the 4 image sequences, only one image sequence is used in the training stage, and the other three are used in the testing stage. Among the 52 individuals, the image samples of 12 persons are used to construct a constrained subspace, and the image samples of the other 40 individuals are used to generate the reference models and to evaluate the recognition performance.

According to the manually marked eye positions, face images are extracted. Each extracted face image is applied first by AST [7] and then resized to 36x36 pixels. In the experiment, the constrained subspace was constructed with 36 training subspaces, r_{no} is set to be 9 and t is set to be 1000.

Form the 40 persons, we randomly selected 35 persons for training, and used all 40 persons for testing. In order to obtain unbiased investigation, we performed the face recognition experiment one hundred times. Finally, the average performance of the one hundred experiments was reported. In all experiments, the parameters are set to $n = 10$, $s_{no} = r_{no} = 9$, $\omega_1 = \omega_2 = 0.5$, $\alpha = 1$ and $\sigma = 0.15$.

The experiment results are evaluated by False Rejection Rate (FRR) and False Acceptance Rate (FAR). Fig. 1 shows the recognition results of the proposed method and those of GDA and CMSM. The recognition rate of the proposed method with no rejection rate is 99.1%, and with a 10% false acceptance rate it is 92.6%. Fig. 2 shows the performance of FAR vs. recognition rate. A decisive recognition means that the current test patterns are recognized to be a specific enrolled person. Let C_no denote the number of correct decisive recognition, and D_no denote the total number of decisive recognition. Then

$$\text{Recognition rate} = \frac{\text{C\_no}}{\text{D\_no}} \times 100\%.$$

The experimental result shows clearly that the proposed GDA+CMSM method is superior to the other two methods.

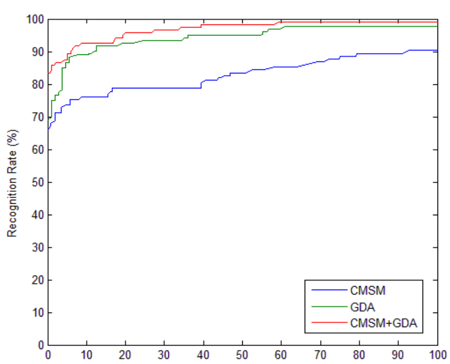


Fig.1 Performance

False Acceptance Rate		0%	10%	20%	30%	100%
Recognition	CMSM	66.3%	76.2%	78.7%	78.7%	90.3%
	GDA	69.1%	89.1%	92.6%	93.4%	97.7%
	CMSM+GDA	83.3%	92.6%	95.8%	96.6%	99.1%

Fig.2 Performance comparison

4. Conclusions

This paper introduces a face recognition method by integrating both single-image and image-sequences matching modules. To diminish the lighting effect, an Anisotropic Smoothing Transform is proposed. Experiments have shown that the proposed method can achieve a very promising recognition accuracy (99.1%) for the famous Banca face database. In the future, we intend to apply the Generalized Discriminant Analysis (GDA) [6] to the single-image recognition and further investigate the recognition performance on some larger face databases.

Acknowledgement

This research is supported in by Industrial Technology & Research Institute under contract 097-B02-009 and by Taiwan NSC under contract NSC972221-E-216-040

References

[1] Matthew A. Turk and Alex P. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*,vol. 3, no. 1, pp. 71–86, 1991.

[2] Li-Fen Chen, Hong-Yuan Mark Liao, Ming-Tat Ko, Ja-Chen Lin, and Gwo-Jong Yu, "A new LDA-based face recognition system which can solve the small sample size problem," *Pattern Recognition*, vol. 33, pp.1713–1726, 2000.

[3] O. Yamaquchi,K. Fukui and K.maeda, "Face Recognition using Temporal Image Sequence," *Pro, Int'l Conf. on Automatic Face and Gesture Recognition*, pp. 318-323, 1998.

[4] K. Fukui, and O. Yamaquchi, "Face Recognition using Multi-viewpoint Patterns for Robot Vision," *Symp. of Robotics Reseach*, 2003.

[5] LinLin Shen and Li Bai, "Gabor Wavelets and Kernel Direct Discriminant Analysis for Face Recognition," *Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04)*.

[6] Baudat and F. Anouar, "Generalized Discriminant Analysis Using a Kernel Approach", *Neural Computation*, Vol. 12, No. 10, pp. 2385-2404.

[7] Yea-Shuan Huang, Fang-Hsuan Cheng and Wei-Cheng Liu, "Face Recognition by Combining Complementary Matching of Single Image and Sequential Images", *11th IAPR Conference on Machine Vision Applications*, Yokohama, Japan, May 2009.

AN IMPROVED ASM-BASED FACIAL FEATURE LOCATING METHOD

[1]Ting-Chia Hsu (許廷嘉), [1]Yea-Shuan Huang (黃雅軒), [1]Shun-Hsu Chuang (莊順旭), [1]Yun-Jiun Wang (王匀駿), [2]Shian Wan (萬象)

[1]Computer Science & Information Engineering Department, Chung-Hua University, Hsinchu, Taiwan
[2]Creativity Laboratory, Industrial Technology Research Institute, Chutung, Hsinchu, Taiwan

ABSTRACT

The active shape model (ASM) has been successfully applied to locate facial feature points. However, the traditional ASM use a grayscale profile as its feature model without considering the different characteristics of landmarks. We cluster all the landmarks into two groups: the corner landmark group and the edge landmark group. In fact, the edge landmark group is further clustered into two subgroups: the facial contour landmark group and the non-facial contour landmark group. Each landmark group has its own specialized feature model. The feature model of the corner landmark group is constructed by using an Adaboost algorithm, the feature model of the facial contour landmark group is a non-symmetrical cross feature model, and the feature model of the non-facial contour landmark group is a symmetrical cross feature model. Experimental results demonstrate that our proposed method can achieve a better performance than the traditional ASM and it also can run in real time.

1. INTRODUCTION

Facial feature extraction is a very popular research field in the recent years which is essential to various facial image analysis such as face recognition, facial expression recognition and facial animation. In general, based on different kinds of information extraction, the technology of facial feature extraction can be divided into two categories. First, local method, which is to detect eye pupils, eye corners, mouth corners, etc. Second, global method, which applies template matching to extract the face contour and the shapes of the eyebrows, eyes, nose and mouth. At present, three kinds of the most commonly used methods are deformable templates (DT) [1], active shape models (ASM) [2][3][4] and active appearance models (AAM) [3][5][6]. Both ASM and AAM are provided by Cootes, they adopted an energy function and by iterative operation, the value of the energy function is optimally decreased. When the energy function reaches its minimum, the best locations of feature points are detected.

In local method, due to the feature models of facial feature points are mutually independent to each other, the detection result is easy to be affected by variation of lighting and poses. In global method, due to it uses multiple feature points and they are probably complementary to each other, it is usually more accurate and robust to motion.

In recent years, ASM has been successfully applied to medical image analysis, such as computed tomography (CT), and it also can be applied to locating facial feature points. However, the accuracy of the facial feature localization is still a problem because face images are much complex than medical images. Therefore, researchers keep on proposing new methods to improve its performance, such as Haar-wavelet ASM [7], SVMBASM [8], ASM based on GA [4], W-ASM [9]. In general these new methods have better accuracy than the original ASM, but they usually need to take much longer computation time.

In this paper, we present a novel feature model to improve the processing accuracy with almost the same processing time. In our method, the facial feature points are divide into two categories, one is edge point and the other is corner point. These two categories have different processing approaches, and they will be explained in Section 3.

This paper is organized as follow. In Section 2, it introduces a classical ASM. Our proposed method using different feature processing models are described in Section 3. Experimental results are given in Section 4, and the conclusion is drawn in Section 5.

2. REVIEW OF THE ACTIVE SHAPE MODEL (ASM)

ASM is one of statistical models, which contains a global shape model and a lot of local feature models. Section 2.1 decides the global shape model; Section 2.2 describes the local feature models and Section 2.3 describes the ASM algorithm.

2.1. The shape model

Supposed there are n facial feature points and each one is located at obvious face contour. The position of these n points can be arranged into a shape vector X, that is

$$X = [\, x_1, y_1, x_2, y_2, \ldots, x_k, y_k, \ldots, x_n, y_n \,]^T$$

where x_k and y_k are the X coordinate and Y coordinate of the k^{th} point respectively.

All the training shapes should be aligned first because we want to obtain the statistic variation of shapes instead of the variation of locations. The ASM alignment procedure is an iterative process to align multiple face shapes which can be summarized as follow:

1. All training sample are normalized according to two eyes positions.
2. Rotate, scale and translate each shape to align with the first shape in the training set.
3. Calculate the mean shape from the aligned shapes.
4. Normalize the mean shape.
5. Realign every shape with the normalized mean shape.
6. If not convergence, return to step 3.

When finishing the alignment procedure, by using the Principal Component Analysis (PCA) operation eigenvectors corresponding to shape variations can be generated. Therefore, a shape model can be represented as:

$$x = \bar{X} + Pb$$

where $\bar{X}$ is the mean shape, $P = [\Phi_1\ \Phi_2\ \ldots\ \Phi_t]$ is the eigenvectors corresponding to the t largest eigenvalues and b is the shape parameter which is the projection coefficiency that X projects onto P. Figure 1 shows the face models of the first three eigenvectors with varying b_i value. Obviously, b_i defines shape variation. In general, the larger b_i is, the more deviation the face shape will be. Usually, b_i is constrained within the range of $\pm 3\sqrt{\lambda_i}$, so that a constructed face shape will not degenerate too much.

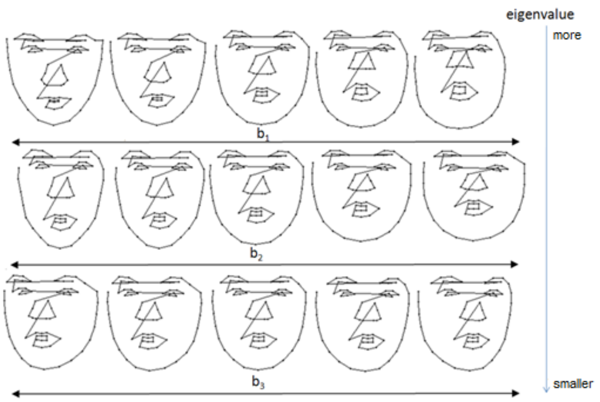


Fig. 1: The variation of the first three parameters of the face model, the horizontal represent the variation value of shape parameter, and the vertical corresponds to the face models derived from different eigenvectors.

2.2. The feature model

In general, we suppose a landmark is located on the strong edge. According to the normal direction of landmark, we can get m pixels on both sides (Fig. 2) of this landmark and each pixel has a gray-level value. So there are in total $2m+1$ gray-level values which form a gray-level profile represented as $g_i = [g_{i0}, g_{i1}, \ldots, g_{i(2m)}]$, where i is the landmark index. In order to capture the frequency information, the profile first derivative dg_j is calculated as

$$dg_i = [g_{i1} - g_{i0}\,, g_{i2} - g_{i1}, \ldots, g_{i(2m)} - g_{i(2m-1)}].$$

In order to lessen the effect of varying image lighting and contrast, the profile is normalized as

$$y_i = \frac{dg_i}{\sum_{k=0}^{2m-1} |dg_{ik}|} \quad \text{where } dg_{ik} = g_{i(k+1)} - g_{ik}.$$

The feature vector y_i is called grayscale profile.

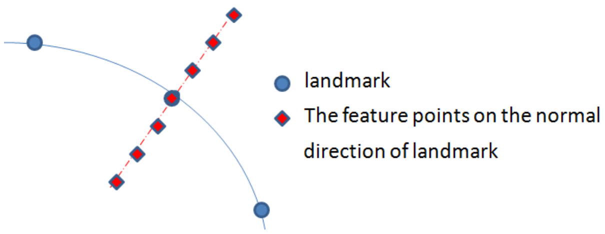


Fig. 2: The selected feature points for constructing the grayscale profile.

2.3. The ASM algorithm

The ASM searching algorithm uses an iteration process to find the best landmarks which can be summarized as follow:

1. Initial the shape parameters b to zero (the mean shape).
2. Generate the shape model point using the $x = \bar{X} + Pb$.
3. Find the best landmark z by using the feature model.
4. Calculate the parameters b' by the following equation
 $b' = P^T(z - \bar{X})$.
5. Restrict parameter b' to be within $\pm 3\sqrt{\lambda_i}$.

If $|b' - b|$ is less than the threshold value, then the matching process is completed; else $b = b'$, then return to step 2.

3. THE PROPOSED METHOD

The traditional ASM uses only the grayscale profile as its feature model which represents the frequency information. However, for certain landmarks such as eye corners and mouth corners their normal directions are difficult to decide and are easy to change significantly. Then the original feature models of these landmarks are very unstable. So that, we had better design a different feature model for them. With this understanding, landmarks are categorized into two groups, the corner landmark group and the edge landmark group. As implied by the name, the corner landmark group contains the landmarks having obvious sharp corner shape, and the edge landmark group contains the landmark having smooth edge shape. In total, the corner landmark group contains 10 landmarks including the inter/outer corners of right/left eyes, the inter/outer corners of right/left eyebrows and the right/left corners of mouth. The rest of landmarks are attributed to the edge landmark group. Fig. 3 shows a few samples of the two categories.

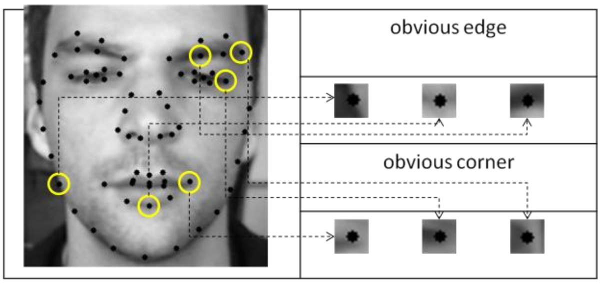


Fig. 3: A few samples of corner landmarks and edge landmarks.

For the edge group, a new feature model called "cross feature model" is proposed which contains three kinds of features: (1) the original grayscale profile with 2m elements, (2) the 2n+1 edge strengths in the tangent direction of landmark, and (3) the edge direction of landmark. Fig. 4 demonstrates the geometry composition of the cross feature model which contains 9 feature points in the normal direction and 3 feature points in the tangent direction. For a landmark i, let y_i denote its grayscale profile, e_i denote its edge strength set, and d_i denote its edge direction. Then the cross feature model C_i can be expressed as $C_i = [y_{i0}, \dots, y_{i(2m-1)}, e_{i0}, \dots, e_{i(2n)}, d_i]$. Obviously, C_i is a (2m+2n+2)-dimensional vector. The computation of y_i is the same as described in Section 2. Suppose the coordinate of e_{ij} is (x, y) and f(x, y) is the gray level of pixel (x,y), then

$$e_{ij} = \sqrt{a_x^2(x,y) + a_y^2(x,y)}$$

where

$$\begin{aligned} a_x(x,y) = {} & [f(x+1,y-1) + 2 \times f(x+1,y) \\ & + f(x+1,y+1)] - [f(x-1,y-1) \\ & + 2 \times f(x-1,y) + f(x-1,y+1)] \end{aligned}$$

and

$$a_y(x,y) = [f(x-1,y+1) + 2 \times f(x,y+1) + f(x+1,y+1)] - [f(x-1,y-1) + 2 \times f(x,y-1) + f(x+1,y-1)].$$

Also, suppose the coordinate of d_i is (x, y), then

$$d_i = tan^{-1}\left(\frac{a_y(x,y)}{a_x(x,y)}\right)$$

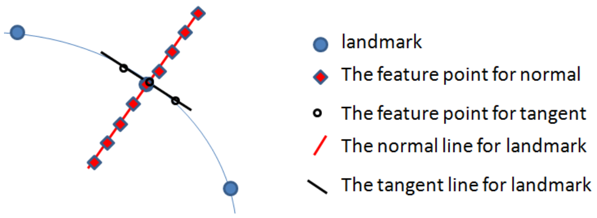
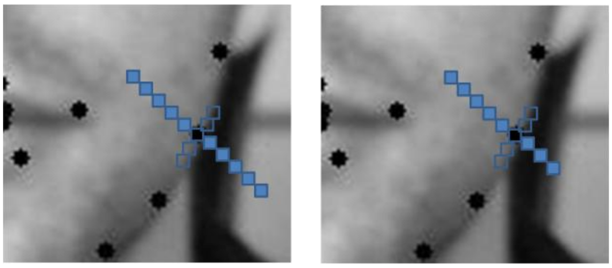


Fig. 4: A diagram to explain the geometric composition of the cross feature model.

In fact, we find the symmetric cross feature is not suitable for the landmarks on facial contour because half of points in the normal direction of such landmark are outside to its face region and their derivative values definitely are unstable. Therefore, for the landmarks located on facial contour, a non-symmetric cross feature model is designed. A non-symmetric feature model is similar to the symmetric feature model and the only difference is in the normal direction of one landmark there are more feature points inside the face region than outside the face region. Fig. 5 shows a non-symmetrical feature model and it has 5 feature points inside the face and 3 feature points outside the face in the normal direction of the selected landmark.



(a) (b)

Fig. 5: (a) the cross feature, (b) the non-symmetric cross feature, it has shorter profile lens, the rectangular of solid is feature point of normal, the rectangular of hollow is feature point of tangent.

For the smooth edge group, the search range of profile has an important role for the processing performance (accuracy and operation speed) because different search ranges will affect the process results considerably. In fact, different landmarks may have different search range for their best performance. With this consideration, landmarks are divided into five clusters according to their locations and structures. The five clusters are the eye cluster, the eyebrow cluster, the nose cluster, the mouth cluster and the facial contour cluster. Because the shape variations of the five clusters are different, we can utilize the shape variation information to determine more appropriate searching ranges so that better landmark matching can be achieved. Therefore, we calculate the standard deviation S_i of the i-th landmark and then the largest standard deviation S_b of each cluster will be used to derive the search range of this cluster. It can be represented as

$$S_b = \max_{i \in b}(S_i)$$

where b is 0 to 4 which corresponds to one specific cluster. Finally, in order to obtain better results, we take twice the length of S_b to indicate the profile search range.

However, some search ranges of the upper eyelid landmarks are overlapped with those of the lower eyelid landmarks. This situation will result in the possibility to mismatch them, i.e. an upper eyelid landmark is mismatched to a lower eyelid landmark or vice versa. In order to avoid this problem, the search range of each eyelid landmark must be constrained. According to the geometry of eye-lid landmarks, the search range of each upper eyelid landmark should not be lower than the middle line between the inner eye corner and the outer eye corner. Similarly, we can define the constraint for the lower eyelid landmark. Fig. 6 illustrates the search range restrictions of the upper and the lower eyelid landmarks.

Fig. 6: Example of the restrictions of the eyelid.

The same issue also happens in the mouth landmarks, but we cannot use the same constraints to restrict their search ranges. This is because people has various expressions and the lower lip may be higher than the connection line of the two mouth corners. Therefore, we intend to use a sorting method to reduce the impact of this issue. Firstly, the mouth landmarks are clustered into three groups: G1, G2 and G3. Fig. 7 shows the clustering results. In the G1 group, when the best matching points of all 4 landmarks have been determined, we can reference the heights of the 4 matching points to order the four landmarks. The order should be the same as the one defined in Fig. 7. The other two clusters G2 and G3 use the same method. Accordingly, the wrong mismatch in mouth will also be decreased.

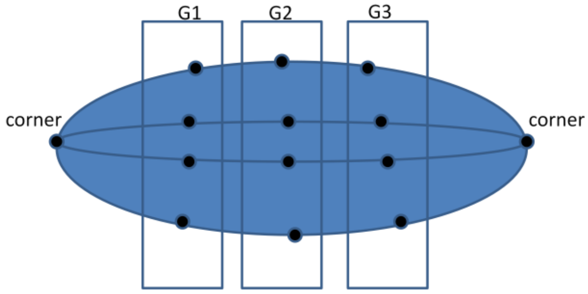


Fig. 7: The clustering results.

Finally, the Mahalanobis distance is used to measure distance between patterns which is computed as

$$f(Y) = \sqrt{(Y - \overline{Y}_i)^T C_i^{-1} (Y - \overline{Y}_i)}$$

where C_i is the covariance matrix of landmark *i*.

We use an Adaboost algorithm to construct a detector for each landmark of the corner group. The Adaboost algorithm has been extensively used for object detection and it often has an outstanding performance. The search range of a corner landmark is $l \times l$ centered by the corresponding landmark location of the reconstructed shape model. If there are multiple detected points among the search range, the point being closest to the corresponding landmark of the reconstructed shape model is taken as the best result. This strategy can avoid abnormal deviation so that it still has a chance to obtain a good matching location in the next iteration.

4. EXPERIMENTAL RESULTS

We use the well known BioID face database as the training database, which contains 1508 face images. In order to increase the training samples, the mirror images will be used, too. In total, there are 3016 face images used in the training stage. The Cohn Kanade database which contains 2132 face images is used for testing. Fig. 8 show some samples of both databases. 67 landmark points are manually labeled for all the images of the two databases.

(a) (b)

Fig. 8: (a)Examples of the BIOID database,
(b) Examples of the Cohn Kanade database.

In order to evaluate the accuracy of our proposed method, the error rate E is defined as

$$E_j = \frac{1}{N}\sum_{i=1}^{N}(\frac{pt_{ij} - ans\_pt_{ij}}{dist_i} * 100\%)$$

where N is the total number of images, pt_{ij} is the matched position of the j^{th} landmark, ans_pt_{ij} is the manually marked position of the j^{th} landmark, and $dist$ is the distance between two eyes.

The hit rates of different corner landmarks will be calculated, and Table 1 shows their individual values and the average hit rate is 96.85%.

Table 1: The hit rates of different corner landmarks

	Hit rate (%)
The left outer eyebrow corner	97.4
The left inter eyebrow corner	94.9
The right outer eyebrow corner	93.9
The right inter eyebrow corner	99.3
The right outer Canthus	96.3
The right inter Canthus	96.6
The left outer Canthus	98.8
The left inter Canthus	98.4
The right mouth corner	94.7
The left Mouth corner	98.2

The overall performance of the modified ASM compared with the traditional ASM is shown in Fig. 9 and some detected results are shown in Fig. 10.

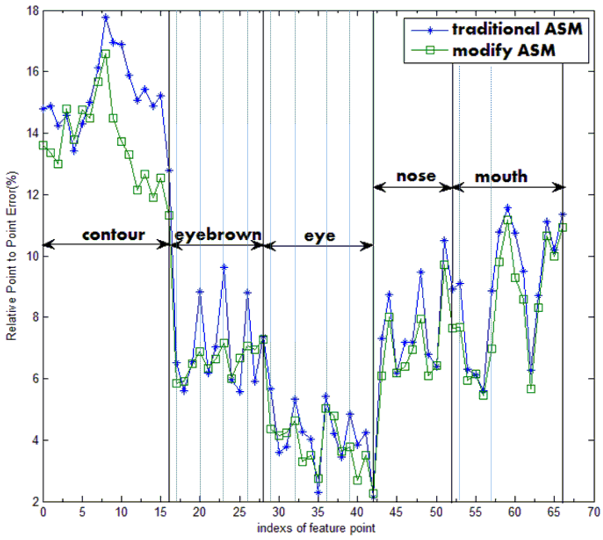


Fig. 9: Errors of each landmark.

From Fig. 9, it obviously shows that the proposed ASM performs better than the traditional ASM on the landmarks of facial contour, nose and mouth. The non-symmetric cross feature model is appropriate to the landmarks of facial contour and the symmetric cross feature model is appropriate to the landmarks of nose and mouth. The corner landmarks of eyebrow, eye and mouth obtain considerably improved performance, such as the *17th*, *20th*, *23th*, *26th*, *29th*, *32th*, *36th*, *39th*, *53th* and *57th* landmark. This means the traditional feature model is not suitable for the corner-like landmarks and they perform unreliable. To process a 640*490 face image, the proposed method takes about 230 ms which can be further speeded up in the near future.

But, our experiments also revealed large error usually is occurred by the change of facial expressions. When the facial expressions are changed significantly among a set of consecutive images, the best matching position probably is not within the profile search region and only the local minimum solution can be obtained. So this problem needs to be further investigated in our future research.

Fig. 10: Some results on the Cohn Kanade database. Top row is the traditional ASM and bottom row is the modifying ASM.

5. CONCLUSION

The traditional ASM use a grayscale profile as the feature model without considering the different characteristics of landmark. We divide all the landmarks into two groups: the corner landmark group and the edge landmark group. From the experimental result, we can achieve the better results than traditional ASM algorithm. In the future work, we will try to define the new searching range and we hope to get better performances.

ACKNOWLEDGEMENT

This research is supported by MOEA under a project 「虛實融合之遊戲科技整合計畫」(project code is 98-EC-17-A-07-07-0708) and by TaiwanNSC under contract NSC972221-E-216-040.

REFERENCES

[1] ZHANG Baizhen and RUAN Qiuqi, "Facial feature extraction using improved deformable templates", The 8th International Conference on Signal Process Nolumn4, page : digital object identifier 10.1109/ICOSP.2006.345927.

[2] T. F. Coots, C. Taylor, D. Cooper, and J. Graham, Active shape models – their training and application, Computer Vision and Image Understanding, 61(1):38-59, January 1995.

[3] T. F. Cootes and C. J. Taylor, Statistical models of appearance for computer vision, Tech. Report, Oct 2001, University of Manchester, http://www.isbe.man.ac.uk/~bim/, Oct 2001.

[4] Kwok-Wai Wan, Kin-Man Lam, Kit-Chong Ng, "An accurate active shape model for facial feature extraction", Pattern Recognition Letters , Volume 26 , Issue 15, November 2005

[5] T. F. Cootes, G. J. Edwards, and C. J. Taylor, Active Appearance Models, in Proc. European Conference on Computer Vision 1998. (H. Burkhardt and B. Neumann Ed.s). Vol. 2, pp. 484~498, Springer, 1998.

[6] T.F. Cootes, G.J. Edwards, C.J. Taylor, Active appearance models, IEEE Transactions on Pattern Analysis and Machine Intelligence 23 (2001) 681–685.

[7] Fei Zuo, Peter H.N. de With, Fast facial feature extraction using a deformable shape model with Haar-wavelet based local texture attributes, Proceedings of IEEE Conference on ICIP, pp. 1425-1428, 2004.

[8] Chunhua Du, Qianq Wu, Jie Yang, Zhenq Wu, SVM based ASM for facial landmarks location, 8^{th} IEEE international Conference on Computer and Information Technology, 2008. CIT 2008.

[9] Feng Jiao, Stan Li, Heung-Yeung Shum, Dale Schuurmans, Face Alignment Using Statistical Models and Wavelet Feature, Proceedings of IEEE Conference on CVPR, pp.321-327,2003.

行政院國家科學委員會補助國內專家學者出席國際學術會議報告

98 年 5 月 27 日

附件三

報告人姓名	黃雅軒	服務機構及職稱	中華大學資工系　副教授
會議 時間 地點	2009.05.20~2009.05.22 日本　橫濱	本會核定補助文號	NSC 97-2221-E-216-040
會議名稱	IAPR Conference on Machine Vision Applications (MVA)		
發表論文題目	Face Recognition by Combining Complementary Matchings of Single Image and Sequential Images		

報告內容應包括下列各項：

一、參加會議經過

- 此次會議有 29 個國家投稿，每篇論文由兩位評審來審稿，最後有 39 篇被接受為 Oral paper, 有 80 篇被接受為 post paper。
- 此會議為 single track，參加者不必急著趕場，可以有較多的時間彼此討論。由於會議專注於電腦視覺的技術和應用，而所有與會人員都是從事於此領域研究，所以有很好經驗交談和學習的機會，
- 在第二天晚宴時，大會主席 Prof. Hideo Saito (Keio University)頒發了 5 篇過去曾在此會議中發表而具有重大影響力的論文，其中四篇來自日本學者，一篇來自韓國學者。每位得獎學者都有發表感言，過程溫馨有趣。
- 此會議安排三個 Invited talks，分別是
  - ✧ (5/20) Large scale image search, Dr. Cordelida Schmid;
  - ✧ (5/21) Focal stack photography: high performance photograph with a conventional camera, Prof. Kyros Kutulakos;
  - ✧ (5/22) Integration of earth observation data: challenge of GEOSS (global earth observation system of systems), Prof. Ryosuke Chibasaki.
- 此會議總共包含 15 sections，每天中午於 13:00~14:30 都有一個 poster section，而在 poster section 之後，也都有一 Invited talk，除此之外還有 9 個 oral presentation

sections，題目為

- ✧ Interaction & virtual reality
- ✧ Motion & multiview
- ✧ Visual surveillance
- ✧ Feature extraction & pattern recognition
- ✧ Human sensing
- ✧ Industrial applications
- ✧ Geographic information systems
- ✧ Machine vision for transportation

二、與會心得

- MVA(Machine Vision Applications)會議除了重視電腦視覺新技術的研發以外，也非常重視應用的開發，例如路標的辨識、腳型的測量、水質的估測和布料的檢查等。這樣性質的研討會頗適合學校老師的參與，不但可以交換研發的心得，還可以看到多方面的應用，刺激老師進行產學合作計畫的動機。
- 此次所頒發的 5 篇過去十年重大貢獻論文獎中大部分得獎者均是日本人，這個現象雖然來自日本人小家子氣的特質外(不願將獎平均分配)，也顯示我們研究的品質需大力的加強，否則對國際社會無法造成實際的幫助。

三、建議

- 本校老師多參與國際性會議，除了介紹研究成果，增加學校的知名度以外，也能快速擴展視野，建立合作管道，對未來研究和教學有很大的幫助。
- 鼓勵學校的研究生參與這種研究與應用結合的國際性會議，讓他們更了解研究的實用價值，以激發學習和研究的熱誠。

四、攜回資料名稱及內容

會議論文集一本和光碟片一片

行政院國家科學委員會補助國內專家學者出席國際學術會議報告

98 年 9 月 16 日

附件三

報告人姓名	黃雅軒	服務機構及職稱	中華大學資工系　副教授
會議 時間 地點	2009.09.12~2009.09.14 日本　京都	本會核定補助文號	NSC 97-2221-E-216-040 NSC 98-2221-E-216-029
會議名稱	International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP2009)		
發表論文題目	Face Recognition Based on Complementary Matching of Single Image and Sequential Images		

五、參加會議經過

- 此次會議有 25 個國家投稿，總共投稿篇數為 410，每篇論文由兩位評審來審稿，最後有 326 篇被接受為 Oral 和 poster 論文。

- 此會議為 multiple tracks，同一時間有 5 個不同主題的 sections 同時舉行，事先要做功課，才能有效的知道要參加的 section。其實，有時候感興趣的論文，可能在同一時段橫跨不同的 sections，則需有效的利用時間，才能得到最好的學習效果。

- 會議中有許多的時間可以彼此討論，有很好經驗交談和學習的機會，甚至談及互訪和未來合作的可能。例如，在第二天晚宴時，大阪市立大學名譽教授 Hiromitsu Hama 談到 12 月份將來台灣參加研討會，他希望有機會拜訪其他大學，我們則表示歡迎他來本校訪問，或許也可邀請他來演講。這個議題將會繼續透過 email 來規劃。

- 會議第二天當我發表論文時，有許多聽眾參加，包含本會議榮譽共同主席 Prof. Joshiaki Shirai 和議程委員會共同主席 Dr. Hitoshi Sakano。當我發表結束時，Dr. Sakano 問了二個問題，當會議結束後，我前去致意，才知道於 2002 年就與 Dr. Sakano 於 ICPR 會議見過面，當時他曾在人臉辨識技術上給我建議。想不到 7 年後我們能在本會議中重逢，雙方都很驚喜，那時會議榮譽共同主席 Prof. Joshiaki Shirai 也加入我們的談話，大家交換了一些研究上的心得，也建立起關係。

- 此會議安排三個 Invited talks，分別是

- ✧ (9/12) The state of the art of 3D video technologies – accurate 3D shape and motion reconstruction, high fidelity visualization, and efficient coding for 3D video, by Prof. Takashi Matsuyama, Kyoto university;
- ✧ (9/13) Data compression by data hiding, by Prof. Hyoung Joong Kim, Korea university;
- ✧ (9/14) Multimodal information fusion in the virtual environment and its applications in produce design, by Prof. Jianrong Tan, Zhejiang university..

- 此會議總共包含 39 sections，其中我參加的 sections 有
  - ✧ Multimedia Signal Processing for Intelligent Applications
  - ✧ Intelligent Surveillance and Pattern Recognition
  - ✧ Advances in Biometrics(I)
  - ✧ Advances in Biometrics(II)
  - ✧ Intelligent Image and Signal Processing
  - ✧ Behavior Analysis and Abnormal Event Detection
  - ✧ Statistical Image Processing and Application
  - ✧ Application of Intelligent Computing to Signal and Image Processing

六、與會心得

- International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP2009)會議包含廣泛的研究議題，都是電腦視覺領域近年來重要的研究領域，藉由與其他學者的交談，可以擴展研究者視野，刺激老師進行產學合作計畫的動機，頗適合學校老師的參與。
- 由於大阪市立大學名譽教授 Hiromitsu Hama 有意願來台參訪，今後將繼續聯絡，促成此事，或許有助於學校在國際化和國際合作等事務的推廣有幫助。

七、建議

- 本校老師多參與國際性會議，除了介紹研究成果，增加學校的知名度以外，也能快速擴展視野，建立合作管道，對未來研究和教學有很大的幫助。
- 鼓勵學校的研究生參與這種研究與應用結合的國際性會議，讓他們更了解研究的

實用價值，以激發學習和研究的熱誠。

八、攜回資料名稱及內容

會議議程一本和光碟片一片