A cascade fusion scheme for gait and cumulative foot pressure image recognition

Shuai Zheng, Kaiqi Huang, Tieniu Tan, Dacheng Tao

National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China
Centre for Quantum Computation & Intelligent Systems, Faculty of Engineering and Information Technology, University of Technology, Sydney, NSW 2007, Australia

Abstract
Cumulative foot pressure images represent the 2D ground reaction force during one gait cycle. Biomedical and forensic studies show that humans can be distinguished by unique limb movement patterns and ground reaction force. Considering continuous gait pose images and corresponding cumulative foot pressure images, this paper presents a cascade fusion scheme to represent the potential connections between them and proposes a two-modality fusion based recognition system. The proposed scheme contains two stages: (1) given cumulative foot pressure images, canonical correlation analysis is employed to retrieve corresponding gait pose image candidates in gallery dataset; (2) pedestrian recognition is achieved via small samples matching between retrieved gait pose images and unlabeled ones. The proposed fusion recognition system is not only insensitive to slight changes of environment and the individual users, but also can be extended to multiple biometrics retrieval. Experimental results are conducted on the CASIA gait–footprint dataset, which contains cumulative foot pressure images and its corresponding gait pose image sequence from 88 subjects. Evaluation results suggest the effectiveness of the proposed scheme compared to other related approaches.

Keywords:
Gait
Foot pressure image
Human recognition

1. Introduction

Recent biomedical and forensic studies reveal that humans can be distinguished by unique walking patterns (e.g., limb movement pattern and ground reaction force). Unique limb movement patterns seem to help to recognize individuals at distance (e.g., gait recognition [2]), while ground reaction force and its variants such as footprints are also separately utilized to identify criminals by human experts [3]. Recent walking pattern based recognition systems are not practical for many reasons due to the single camera sensor. For example, viewpoint problems in gait recognition under constraints could be avoided if we use Kinect [4]. The cumulative foot pressure image is a cumulative image-type record of ground reaction force change during one gait cycle [5]. It provides richer information to identify different walking patterns, compared to the existing 1D ground reaction force or simple 2D footprint pictures. The cumulative foot pressure image has been applied in biomedical assistant, forensic investigation, sports assistant training and custom shoes [5]. In order to address the problems in existing walking pattern recognition systems, we propose to develop a computational correlation model for cumulative foot pressure images and gait.

As far as we know, there are a few previous works proposed to develop the computational correlation model for cumulative foot pressure image and gait, although a lot of works have been proposed to study gait recognition [6–8] or the cumulative foot pressure images [5]. Multimodal biometric system [1] have been proposed to combine evidences from different sources. These sources might simultaneously come from various sensors [9], different classification algorithms [10], multiple instances for the evidence or directly from diverse biometric traits [11]. Naive feature combinations may not always improve the performance, since some components in different sources may not be complementary. Zhang et al. found that the performance of human recognition using multiple sources could be improved by reducing the redundant classes in the gallery dataset [12]. Specifically, in order to evaluate the computational correlation model we obtain, we also develop a human recognition system using gait pose images and corresponding cumulative foot pressure images without these limitations using a cascade fusion scheme.

The proposed study is necessary since it not only provides a computational correlation model but also a solution for entrance control applications. For example, in jailhouse security system or suspect identification, cumulative foot pressure images and gait pose images of the same individual can be captured at different times. Hence, cumulative foot pressure images and gait pose

* Corresponding author. Tel.: +86 10 82610278; fax: +86 10 62551993.
E-mail addresses: szheng@nlpr.ia.ac.cn, kylezheng04@gmail.com (S. Zheng), kqhuang@nlpr.ia.ac.cn (K. Huang), tnt@nlpr.ia.ac.cn (T. Tan), dacheng.tao@uts.edu.au (D. Tao).

0031-3203/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2012.03.008
images can be used to identify escaped criminal or investigate suspects noninvasively. Furthermore, the proposed cascade multimodal fusion scheme can also be employed to develop a cross-modal fusion approach in gait recognition. The state-of-the-art approaches in gait recognition can be divided into model-based and model-free approaches.

The remainder of this paper is organized as follows. Section 2 presents related work about gait, cumulative foot pressure and multiple evidence fusion schemes for pedestrian recognition. Section 3 illustrates the proposed cascade fusion scheme. In Section 4, feature representation for gait pose images and cumulative foot pressure images are presented. Section 5 describes the fusion based recognition system. Section 6 introduces the dataset. Section 7 reports the experiments. We draw conclusions in Section 8.

2. Related works

Gait recognition is potentially useful for personal identification [14]. It is quite attractive for identification purposes since its advantages are that it is completely unobtrusive, and does not involve any subject cooperation or contact. The state-of-the-art approaches in gait recognition can be divided into model-based and model-free approaches.

Model-based approaches tend to recover the underlying mathematical construction of gait with a structure motion model. The mean shapes of gait silhouettes are modeled by Wang et al. via employing procrustes analysis [15]. Bouchrika and Nixon extract crucial feature descriptions from human joints by developing a motion-based model using elliptic Fourier descriptors [16]. However, the performance of the approaches suffers from poor localization of the torso and difficult extraction of underlying models from gait sequences.

The other kind of approach is model-free one. One kind of model-free approach preserves temporal information in recognition and training states. Hidden Markov models (HMMs) are utilized to achieve gait recognition [17]. Principal component analysis (PCA) [18,19] is employed to extract statistical spatial-temporal feature descriptors of gait [20]. In this kind of approach, large-scale training samples are required for probabilistic temporal modeling approaches to obtain a good performance. Hence, the disadvantage for the approach is the high computational complexity of sequence matching during recognition and the high storage requirement of the dataset. Another kind of model-free approach converts a sequence of images into a single template. Gait recognition by averaging all the silhouettes is presented by Liu et al. [21]. Han proposed a gait energy image (GEI) to construct real and synthetic gait templates [22]. The recognition performance may degrade since the temporal information in gait sequences are discarded. Wang et al. developed a spatial-temporal walking template called chrono-gait image (CGI) to encode the temporal information via color mapping to improve the recognition rates [23]. The main drawback of these approaches is that they easily suffer from slight changes of environment such as illumination variation in probe and gallery data collections or crowded scenarios. Besides, traditional motion-based gait representation is not practical and stable in more wide application scenarios such as internet videos or image sequences from IP camera.

Recent works on action recognition started to introduce some feature descriptors like histograms of oriented gradients (HOG) to represent several action key poses [24]. Such kinds of feature descriptors have also helped to achieve state-of-the-art performance in object detection and object recognition [25]. In these tasks, they are proved to overcome environmental challenges and be able to represent objects without background priority or
motion-based segmentation. Inspired by their success, we propose to use combination feature descriptors obtained from several continuous gait pose images to represent individual gait.

Footprints are an important identification evidence for forensic investigation purposes. Although it has been applied since ancient times, only recently Kennedy first proved the uniqueness of barefoot prints and their use as a possible means of identification [26]. Previous works can be divided into two types, one is ground reaction force and another is a still footprint image. Ground reaction force is a 1D data signal record of walking. Moustakidis proposed a subject recognition system based on ground reaction force measurement [27]. Cattin developed a general PCA fusion based biometric system using both gait and ground reaction force [7]. However, this method is neither robust to slight noise nor able to distinguish different walking manners between different pedestrians. Besides, the strictly controlled data collection environment limits its potential applications. The second method is utilizing still footprint image to achieve pedestrian recognition. Nakajima proposed person recognition using normalized pairs of raw barefoot prints [28]. Uhl developed a footprint-based biometrics verification system using scanned barefoot images [29]. The method failed in recognizing the same individual when users wore different shoes. Cumulative foot pressure images contain cumulative spatial and temporal force information during one gait cycle, which may help to handle the difficulties in recognizing individuals wearing different shoes. Previous work shows that feature descriptors based on hierarchical models are invariant to different shoes [30]. Inspired by the success of hierarchical models on translation-invariant object recognition datasets, we propose a Gaussian mixture model (GMM) based on cumulative foot pressure images.

Existing multimodal biometric systems are proposed by combining evidence from different sources [7,31]. Different combination approaches can be divided into three types: the feature level, matching score level and decision level. Cattin utilized Bayesian decision theory to fuse ground reaction force and gait [7]. This loses much correlation information. Zhang et al. proposed to fuse evidence at the feature level [31]. However, this is not practical since the multiple modalities may be incompatible and direct concatenating feature vectors may suffer from the “curse of dimensionality” problem. He et al. investigated the performance of various score level fusion approaches in multimodal biometric systems [32]. But these approaches do not address the issue of fusing evidence obtained at different times. Motivated by multimodal document cross retrieval systems [13], we propose a cascade fusion scheme to fuse gait and cumulative foot pressure images both at the feature level and score matching level.

3. Feature representation

In this section, we attempt to achieve gait representation from still image sequences and a cumulative foot pressure image representation which is translation invariant.

3.1. Gait representation

A very basic assumption in all gait recognition research is that all walking sequences from the same person follow a similar walking pattern, where the walking pattern involves the moving range of the limbs. However, there are many walking cycles repeated in one walking image sequence. To estimate the walking period and separate a single walking cycle from one walking image sequence, we compute the movement of the limbs as the frame changes. We firstly employ Felzenszwalb’s detector [33] to extract the bounding box for the pedestrian in each frame. Then we compute the width change of bounding box as the frame changes. This is similar to the method of Sarkar et al. [21].

We compute the gait representation as the concatenated probability of boundary based histogram of oriented gradients (pbHOG) descriptors based on the normalized cropped walking poses for one walking cycle. We extract the HOG descriptors based on the probability of boundary (Pb) operator [34] responses. As Fig. 2 illustrates, the Pb operator suppresses small noise in the image. Hence, pbHOG captures more salient walking pose details.

3.2. Translation-invariant representation for cumulative foot pressure image

Our goal is to learn a feature representation model for cumulative foot pressure images which preserves translation-invariance. Recent work [35] has been proposed to address the shoe-invariant cumulative foot pressure image recognition. In contrast, we address the problem of slight changes in barefoot pressure images. Current low-level descriptors are proved to be invariant to minor translation and effective in many object recognition applications. However, they are not practical for representing cumulative foot pressure image since they lose a lot of structure information (in terms of object parts) which is crucial for shoes-invariant cumulative foot pressure image recognition. Previous works [12] prove

![Fig. 2. Comparison of different gait representations and overview of concatenated pbHOG gait pose image representation.](image)
that hierarchical model will bring advantages of translation-invariant power. Hence, we propose to model cumulative foot pressure image with hierarchical Gaussian mixture models, as Fig. 3 illustrates.

Suppose $z$ denotes a $p$-dimensional feature vector (SIFT descriptor) or COP (center of pressure curve) [27, 5] from cumulative foot pressure image $l$. We model $z$ by GMM model:

$$p(z|\theta) = \sum_{k=1}^{M} w_k^l N(z; \mu_k, \Sigma_k),$$

where $M$ denotes the total number of Gaussian components and $(\mu_k, \Sigma_k)$ are the weight, mean and covariance matrix of the $k$th Gaussian component, respectively. For computational efficiency, the covariance matrices $\Sigma_k$ are restricted to be a diagonal matrix. The number of model parameter $\theta = (\mu_k, \Sigma_k)$ increases with respect to $N$. $N$ is the number of images, $p(\theta)$ is the distribution of the parameters. Following [36], the prior mean vector, prior weights, and covariance matrix are estimated by fitting a global GMM. The other parameters are learned by maximum a posteriori (MAP) loss:

$$\max \ln [p(z|\theta)] + \ln p(\theta).$$

Center of pressure histograms in the cumulative foot pressure image and dense SIFT feature descriptors are separately extracted and modeled with GMMs. We achieved a CFPI representation via concatenating these descriptor vectors. We represent the cumulative foot pressure image with the parameters of the GMM. Hence, following the suggestion in [36], we represent cumulative foot pressure image as

$$x = [\sqrt{w_1 \Sigma_1^{-1/2} \mu_1}, \ldots, \sqrt{w_M \Sigma_M^{-1/2} \mu_M}].$$

4. Cascade fusion scheme

In this section, we present a cascade fusion scheme using canonical correlation analysis and its variant regression algorithm.

4.1. Feature selection using canonical correlation analysis

In order to identify shared correlated structure among variables from two evidence sources, Canonical correlation analysis (CCA) [37] is employed as a feature selection method firstly. The algorithm estimates two basis vectors so that, after linear projection, the correlation between the two classes is mutually maximized. Given two sets of sample vectors $S = (r_1, u_1), \ldots, (r_n, u_n)$, and their projection matrices $w_r$ and $w_u$, CCA mutually maximizes the object function:

$$\rho = \max_{w_r, w_u} \frac{w_r^T C_{rr} w_r}{\sqrt{w_r^T C_{rr} w_r} \sqrt{w_u^T C_{uu} w_u},}$$

where $C_{rr}$ and $C_{uu}$ are within-sets covariance matrices and $C_{ru}$ is between-sets covariance matrices. A closed form solution can be computed by solving a generalized eigenvalue problem. Large problems can be solved efficiently using predictive low-rank decomposition with partial Gram–Schmidt orthogonaliza-

4.2. Correlated scores computation via CCA regression

In order to evaluate the correlation scores between the data from two evidences, CCA regression is adopted again. Given probe feature vectors $X : x_1, \ldots, x_n$ and gallery feature vectors $Y : y_1, \ldots, y_n$, a mapping pair $W_x, W_y$ is learned via CCA regression:

$$\arg \max_{W_x, W_y} \frac{E[W_x^T X Y^T W_y]}{\sqrt{E[W_x^T X X^T W_x]} \sqrt{E[W_y^T Y Y^T W_y]}}.$$

Based on the mapping pair, we can compute projections as follows:

$$x' = W_x^T x,$n

$$y' = W_y^T y.$n

Correlated scores $s(x_i', y_j')$ are computed between each gallery and probe via CCA regression:

$$s(x_i', y_j') = \frac{x_i' y_j'}{\|x_i'\| \|y_j'\|}.$$

Based on the scores, gallery samples that have high correlated scores are retrieved as candidates for further pedestrian recognition.

5. Cascade fusion scheme for two-modality based pedestrian recognition

Based on the proposed cascade fusion scheme, we achieve pedestrian recognition using gait and cumulative foot pressure images. As Fig. 4 illustrates, there are two inputs of the system: cumulative foot pressure images and corresponding continuous gait pose images. The two evidence sources can be simultaneous or not, but need to be collected from the same individual during
labeled dataset construction. Following Section 3, the concatenated HOG descriptors are proposed to represent a given gait while a super-vector using GMM parameters is employed to preserve the characteristics of given cumulative foot pressure images. In the cascade fusion scheme, cumulative foot pressure images are firstly employed to prune irrelevant parts of the search space in the labeled gait dataset and predict the correlation similarity scores between the given cumulative foot pressure images and labeled gait image sequences. Then, the pedestrian recognition process is achieved by matching the given gait image sequences with a small number of labeled ones, using a trained linear SVM classifier.

6. Dataset

As far as we know, this dataset is the first publicly accessible dataset containing both cumulative foot pressure images and corresponding gait together. This dataset consists of 3496 gait pose images and 2658 cumulative foot pressure images. The distributions of subjects in some basic attributes are presented in Fig. 5. The data were collected from 88 pedestrians, 20 female and 68 male, in an indoor environment. Experimental factors like illumination, background and clothes are not under strict control, as Fig. 6 illustrates.

7. Experiments

The purpose of the proposed system is to develop a computational evaluation framework for studying the relations between gait and cumulative foot pressure images. Although the applications of the study presented in this paper are not limited to recognition, we evaluate the performance of the proposed approach following recognition evaluation criteria. To evaluate the effectiveness of proposed cascade framework in recognition, we choose three other comparison approaches including a single source approach (only gait feature descriptor), naive baseline fusion approach, and CCA based fusion approach [10].

In experiments, we randomly divided the dataset into two subsets: a training set of 440 groups of data containing gait pose images and correlated cumulative foot pressure images and the testing set containing the other 434 groups of data.
We evaluate the recognition performance using accuracy. To get the best results, we compare the different potential features and feature selection approaches considering the performance of retrieval, which often uses precision and recall as evaluation criteria. Suppose the number of retrieved gait pose images is $R_g$, and the number of the retrieved gait pose images which is correlated with given cumulative foot pressure is $R_{cg}$. The precision is defined as

$$\text{precision} = \frac{R_{cg}}{R_g}.$$ 

Recall is the fraction of the gait pose images which are correlated to the correct retrieved ones. Assume that the number of correlated gait pose images is $R_c$. Then, the recall is defined as

$$\text{recall} = \frac{R_{cg}}{R_c}.$$ 

### 7.1. Feature and feature selection

To choose the features and feature selection approach, we perform a retrieval evaluation. During the training stage, we learn the feature subspace projection matrix. During testing, we utilize the learned projection matrix to reduce the dimensionality of the pbHOG and GMM feature vectors. Then our proposed CCA-based cascade scheme is employed to assign correlation scores between gait pose images in dataset and cumulative foot pressure image queries.

As Fig. 7 illustrates, the pbHOG feature descriptor and PCA–CCA feature selection are the best combination. Considering the comparison of different feature descriptors, pbHOG performs best since the probability of boundary largely reduces noise response in still images which helps to reduce the within class divergence. Considering different feature selection approaches, PCA and CCA combination performs best because of the natural advantages of PCA and CCA. PCA captures the principle components and ensures that the feature vectors from the two evidence sources are not illposed, while CCA maximizes the correlations between cumulative foot pressure images and gait pose for the same person.

The precision–recall curve is presented only at the recall level at 10.67% because it is difficult to visualize the precision–recall curve perfectly since we only have 88 subjects. Hence, rather than giving the whole precision–recall curve, we give comparison results of precision values under two different recalls. Recall is the fraction of relevant instances that are retrieved. The precision

---

**Fig. 6.** The sample of dataset containing gait pose and cumulative foot pressure images.

**Fig. 7.** Precision of GPIs pruning using CFPIs images as queries, when recall value is 10.67%. The experimental results show that PCA + CCA feature selection method with pbHOG GPIs representation outperforms others.
drops significantly when the recall increases. The corresponding precision value is too small when recall is larger than 10.67%. In this experiment, we find that the system achieves best performance when we pick up the recall at 10.67%.

7.2. Fusion scheme for pedestrian recognition

To evaluate the performance of the proposed fusion approach, we perform a recognition performance evaluation. During the training stage, feature selection projection matrices and linear SVM classifier are trained on the gallery dataset. During the testing stage, there are two inputs: one is cumulative foot pressure images, another is gait pose images. Given probe data in the form of cumulative foot pressure images and gait pose images, the proposed approach will feed back the corresponding label according to the matching process between the probe and gallery data. Specifically, the process in our proposed cascade scheme for pedestrian recognition is presented as follows: First, unlabeled cumulative foot pressure images are used as a probe data to retrieve correlated labeled gait pose images in gallery dataset. Then, the correlation score helps to prune much irrelevant labeled gait pose images and obtain a small number of labeled gallery dataset. Second, unlabeled gait pose images are sent to the classifier via matching with the small number of labeled gallery dataset.

In our experiments, the penalty parameter C in the linear SVMs are all set as 10. In the cascade scheme, the threshold is set to 10, which means we choose only the top 10 correlated scores of labeled gait pose images for the gait matching process. We choose this threshold because the recognition performance of proposed system achieves the best results under this setting.

In Fig. 8, we choose the result of PCA feature selection based gait recognition to compare the fusion scheme with human recognition using only single source.

In the other approaches illustrated in Fig. 8, we compare different fusion schemes while fixing the feature representation and feature selection approach. For the feature representation and feature selection approach, we use phHOG for gait representation and the GMM representation for cumulative foot pressure images while we use PCA–CCA as feature selection. Naive fusion means combining the dimension reduced feature vectors from gait and cumulative foot pressure image directly together. The benchmark CCA based fusion method is the same as illustrated in [10]. Cascade fusion is the method developed in this paper.

Fig. 8 shows the performance result of different pedestrian recognition schemes. The correlated-model-based cascade two-modality fusion scheme outperforms the other simple concatenated ones. This is not a surprising result. The correlated model helps to reduce the number of labeled GPIs in the gallery dataset and prune the irrelevant GPIs. Hence, it casts the original difficult multi-class matching problem into a small-class or even binary class matching problem. Further, the proposed cascade method outperforms the concatenated-based fusion scheme. This is because the proposed method utilizes the correlation between the two modal data in correlation common space, while the concatenated-based ones ignore this information.

8. Conclusions and future work

This paper reveals a new problem on how to reveal the walking behavior pattern and demonstrate a possible way to exploit the correlation among foot pressure, footprint and walking motion patterns for recognition. Specifically, we presented a work that allows to use different walking patterns to recognize pedestrian without camera setting. The gait pose and cumulative foot pressure are two types of correlated person walking behavior patterns. In order to study the potential connections between the two walking behavior patterns, we establish a standard database, develop a pedestrian recognition system with the cascade fusion scheme. The cascade scheme is proved to be effective to extract the correlated parts between the two modalities, according to the evaluation on the pedestrian recognition performance.

The proposed cascade fusion scheme could be used in designing a cross-biometric searching system which allows to retrieve pedestrian or any biometric pattern with other query biometric patterns. There are also some applications of interest without camera settings, especially for criminal investigation and other related specific applications. For example, it will be more convenient to prune the suspect criminal or others via comparing among footprint, cumulative foot pressure images and gait pose images.

The drawback of this paper is to assume that all subjects are not wearing shoes, which is not very convenient for normal use. Wearing different shoes would generate different footprint shapes which might cause the recognition failure. In this case, future work should be conducted to address the deformable invariant cumulative foot pressure image recognition. Further, the future work will also focus on studying the correlation between footprint and cumulative foot pressure images, so that the computational model can help forensic investigation, which often collects a lot of footprint data from a crime scene.

Acknowledgment

This work is supported by National Basic Research Program of China (2012CB316302) and National Science Foundation of China (Nos. 60875021 and 61175007). It also in part supported by ARC discovery project with Project No. DP-120103730. We also appreciate the help and suggestions of Dr. Tianyi Zhou, Prof. Liming Shi, Prof. Liang Wang, and Dr. Ran He.

References


Shuai Zheng received the B.Eng. from Beijing Institute of Technology in 2008. He is now a master student from National Laboratory of Pattern Recognition, at the Institute of Automation, Chinese Academy of Sciences. His current research interests include computer vision and pattern recognition, especially on image representation, image classification, object detection and gait recognition. He is a student member of the IEEE and the IEEE Computer Society.

Kaiqi Huang received his B.Sc. and M.Sc. from Nanjing University of Science and Technology, China, and obtained his Ph.D. degree from Southeast University. He has worked at the National Laboratory of Pattern Recognition (NLPR), the Institute of Automation, and the Chinese Academy of Science (CASIA) in China and he has been an associate professor at NLPR since 2005. He is a Senior member of the IEEE. He is an executive team member of the IEEE SMC Cognitive Computing Committee as well as Associate Editor of the International Journal of Image and Graphics (IJIG) and Electronic Letters on Computer Vision and Image Analysis (ELCVA) as well as Guest Editor of a Signal Processing special issue on Security. His current research interests include visual surveillance, digital image processing, pattern recognition and biological based vision. He has published over 80 papers in important international journals and conferences such as IEEE TIPAMI, T-IP, T-SMC-B, TCVT, Pattern Recognition (PR), Computer Vision and Image Understanding (CVIU), ECCV, CVPR, ICIP, IJCB. He received the Best Student Paper Awards from ACPRT10 and was a prize winner in the detection task in both PASCAL VOC10 and PASCAL VOC11. He received the honorable mention prize for the classification task in PASCAL VOC11.

Tieniu Tan received the B.Sc. degree in Electronic Engineering from Xi’an Jiao tong University, China, in 1984 and the M.Sc., D.I.C. and Ph.D. degrees in Electronic Engineering from Imperial College of Science, Technology and Medicine, London, UK, in 1986, 1986, and 1989, respectively. He joined the Computational Vision Group, Department of Computer Science, The University of Reading, England, in October 1989, where he worked as research fellow, senior research fellow, and lecturer. In January 1998, he returned to China to join the National Laboratory of Pattern Recognition, the Institute of Automation of the Chinese Academy of Sciences, Beijing. He is currently a professor and director of the National Laboratory of Pattern Recognition as well as president of the Institute of Automation. He has published widely on image processing, computer vision and pattern recognition. His current research interests include speech and image processing, machine and computer vision, pattern recognition, multimedia, and robotics. Dr. Tan serves as referee for many major national and international journals and conferences. He is an associate editor of Pattern Recognition and of the IEEE Transactions on Pattern Analysis and Machine Intelligence and the Asia editor of Image and Vision Computing. He was an elected member of the Executive Committee of the British Machine Vision Association and Society for Pattern Recognition (1996–1997) and is a founding co-chair of the IEEE International Workshop on Visual Surveillance.

Dacheng Tao is currently a professor with the Centre for Quantum Computation and Information Systems and the Faculty of Engineering and Information Technology in the University of Technology, Sydney. He mainly applies Statistics and Mathematics for data analysis problems in data mining, computer vision, machine learning, multimedia, and video surveillance. He has authored and co-authored more than 100 scientific articles at top venues including IEEE T-PAMI, T-KDE, T-IP, and NIPS, with best paper awards.