Journal of Zhejiang University SCIENCE A
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

General moving objects recognition method based on graph embedding dimension reduction algorithm*

Yi ZHANG†, Jie YANG, Kun LIU

(*Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, Shanghai 200240, China*)

†E-mail: zypk@sohu.com

Received June 25, 2008; Revision accepted Oct. 6, 2008; Crosschecked May 19, 2009

Abstract: Effective and robust recognition and tracking of objects are the key problems in visual surveillance systems. Most existing object recognition methods were designed with particular objects in mind. This study presents a general moving objects recognition method using global features of targets. Targets are extracted with an adaptive Gaussian mixture model and their silhouette images are captured and unified. A new objects silhouette database is built to provide abundant samples to train the subspace feature. This database is more convincing than the previous ones. A more effective dimension reduction method based on graph embedding is used to obtain the projection eigenvector. In our experiments, we show the effective performance of our method in addressing the moving objects recognition problem and its superiority compared with the previous methods.

Key words: Moving objects recognition, Adaptive Gaussian mixture model, Principal component analysis, Linear discriminant analysis, Marginal Fisher analysis

doi:10.1631/jzus.A0820489 **Document code:** A **CLC number:** TP317.4

INTRODUCTION

Moving objects may be pedestrians, automobiles, animals, military targets, etc. It is necessary to recognize the object and then decide how to analyze its trajectory or behavior. Therefore, moving objects classification or recognition is the basic and key problem in visual surveillance systems. The main purpose of visual surveillance is to detect then track certain targets and more generally to analyze their trajectories and behaviors in certain areas. The Carnegie Mellon University (CMU) system (Collins *et al*., 2000) can monitor activities over a large area using multiple cameras that are connected to a network. It can detect and track multiple persons and vehicles within cluttered scenes and monitor their activities over long periods of time. The real-time visual surveillance system W4 (Haritaoglu *et al*., 2000) employs a combination of shape analysis and tracking. It constructs models of people's appearances to detect and track groups of people while monitoring their behaviors, even in the presence of occlusions and in outdoor environments. Duque *et al*.(2007) presented a novel behavior analysis system called dynamic oriented graph (DOG). It is used to detect and predict abnormal behaviors, using real-time unsupervised learning. The DOG method characterizes observed actions by means of a structure of unidirectional connected nodes. Each one defines a region in the hyperspace of attributes measured from the observed moving objects and is assigned a probability to generate an abnormal behavior. These techniques are widely used in many situations, such as access control in special areas, person-specific identification in certain areas, crowd flux statistics and congestion analysis, anomaly detection and alarming, and interactive surveillance using multiple cameras (Hu and Tan, 2004).

Related works

Our study focuses on moving objects feature extraction and recognition methods. Feature extrac-

* Project (No. 60805001) partially supported by the National Natural Science Foundation of China

tion and recognition methods are the main areas of research related to the moving objects recognition problem. Different models have been established to define moving objects.

Lipton *et al.*(1998) used 'dispersedness' which is defined as the ratio of the perimeter to the area of moving objects as a classification feature. Humans and vehicles are distinguished from background clutter and tracked by a combination of temporal differencing and template matching. Liu and Su (2000) used two feature parameters of the human body extracted from the horizontal and vertical projection histogram of a silhouette pattern. Rivlin and Rudzsky (2002) developed a best inter-class separation method based on geometric properties of the fitted ellipse and the 'star skeleton'. Collins *et al.*(2000) also used image blob dispersedness, image blob area and apparent aspect ratio of the blob bounding box, etc., as key features. Cutler and Davis (2000) described a similarity-based technique to detect and analyze periodic motion. By tracking a particular object, its self-similarity is computed as it evolves over time. As we know, for periodic motion, its self-similarity measure is also periodic. Therefore, time-frequency analysis is applied to detect and characterize the periodic motion, and tracking and classification of moving objects are implemented using periodicity.

The feature extracting algorithms listed above have been developed specifically for moving objects such as humans or automobiles. When other objects are introduced to be recognized or tracked, a new feature needs to be incorporated into the algorithm. Using a global feature such as the principal component analysis (PCA) coefficient for an object can be a more general approach. Boosting based detection methods have been successfully used for robust detection of faces and pedestrians (Javed and Ali, 2005). The features used for classification are derived from PCA of the appearance templates of the training examples. A detection and classification system for moving objects has been described using range images via a Time of Flight camera and a stereo vision system (Ghobadi and Hartmann, 2006). The range images are used as the input data. Two different classifiers have been trained based on heuristic and PCA features for each image set. This is a classical linear dimensionality reduction approach. Munder and Gavrila (2006) presented an in-depth comparative experimental study on pedestrian classification. Global versus local and adaptive versus non-adaptive features, as exemplified by PCA coefficients, Haar wavelets and local receptive fields (LRFs) were compared.

Using a general global feature to recognize a certain object among various moving objects is a useful approach, for there is no need to define new features when we want to recognize a new object. A good global feature needs to be able to preserve the intrinsic features of several different moving objects. Many linear dimensionality reduction algorithms and dimensionality reduction algorithms based on manifold learning aim to discover the underlying structure of the data. A dimension reduction algorithm, marginal Fisher analysis (MFA) based on graph embedding, was designed by Yan *et al.*(2007) to find the intraclass compactness and the interclass separability of different datasets. In our experiments, we found that for a moving objects classification problem it is more suitable than other dimensionality reduction algorithms. Its principles and advantages will be analyzed in Section 4.

Our research roadmap

In our work, we firstly extract all possible moving objects in the scene with adaptive Gaussian mixture model (AGMM). These objects are then unified to a certain form to build our training database. The training database is dimensionality reduced and trained using MFA to obtain a projection direction. When new objects enter our scene, they will also be unified and dimensionality reduced, and then classified into corresponding categories using the *K*-nearest neighbors (KNN) classifier. The flowchart is illustrated in Fig.1. We have shown in our experiments, that if the training database contains large numbers of items, we can obtain reliable and consistent recognition results. This method has the advantage that it is not limited to certain classes of objects and the objects can be rigid or non-rigid. If an interesting object appears, it will be tracked and analyzed. We use a mean shift tracking algorithm to solve an object's partial occlusion or conglutination problem.

In Section 2, we introduce the background subtraction algorithm used for segmentation of moving objects. In Section 3, different subspace feature extracting algorithms are analyzed and compared.

Section 4 describes the establishment process of our moving objects training and testing database. The effectiveness of different dimensionality reduction methods and their recognition performances are compared. Section 5 concludes the paper.

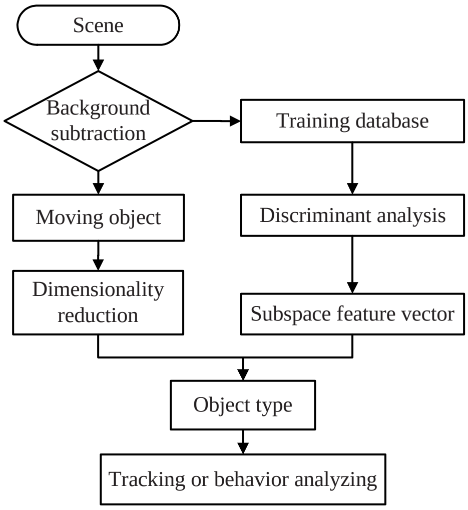

Fig.1 System flowchart of moving objects recognition

BACKGROUND SUBTRACTION

To establish a robust visual surveillance system, the background subtraction algorithm should be adaptable and flexible enough to deal with variations in lighting, moving scene clutter, multiple moving objects and other arbitrary changes to the observed scene. The AGMM (Stauffer and Grimson, 1999) can be used to determine which Gaussian distribution is most likely to result from the background process. An on-line approximation is used to update the model. It has proved to be a stable and efficient background subtraction method (El Baf *et al.*, 2007) and is used in our system to extract moving objects and establish the sample database.

At time t, the historical value of a particular pixel (x_i, y_i) is defined as

$$\{\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_t\}=\{I(x_i, y_i, i), 1\le i\le t\}. \quad (1)$$

Then $\{\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_t\}$ is modeled by a mixture of K Gaussian distributions. The probability of observing the current pixel value is

$$P(\boldsymbol{X}_t)=\sum_{i=1}^{K}\omega_{i,t}\times\eta(\boldsymbol{X}_t,\mu_{i,t},\boldsymbol{\Sigma}_{i,t}), \quad (2)$$

where K is the number of distributions, determined by the available memory and computational power and usually set from 3 to 5, $\omega_{i,t}$ is an estimate of the weight of the ith Gaussian in the mixture at time t, $\mu_{i,t}$ is the mean value of the ith Gaussian in the mixture, $\boldsymbol{\Sigma}_{i,t}$ is the covariance matrix of the ith Gaussian in the mixture, and η is a Gaussian probability density function:

$$\eta(\boldsymbol{X}_t,\boldsymbol{\mu},\boldsymbol{\Sigma})=\frac{1}{(2\pi)^{t/2}\|\boldsymbol{\Sigma}\|^{1/2}}\mathrm{e}^{-\frac{1}{2}(\boldsymbol{X}_t-\boldsymbol{\mu}_t)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}_t-\boldsymbol{\mu}_t)}. \quad (3)$$

Thus, the distribution of recently observed values of each pixel in the video is modeled by a mixture of Gaussians. When a new pixel value comes, it is checked against the existing K Gaussian distributions, until a match is found. A match is defined as a pixel value within 2.5 standard deviations of one distribution. If the current pixel value matches none of the K distributions, the least probable distribution is replaced with a distribution whose mean value is set with the current pixel value, meanwhile an initially high variance and low prior weight are defined.

The prior weights $\omega_{k,t}$ of the K distributions at time t are adjusted as follows:

$$\omega_{k,t}=(1-\alpha)\omega_{k,t-1}+\alpha M_{k,t}, \quad (4)$$

where α is the learning rate and $M_{k,t}$ is 1 for the model that matched and 0 for the remaining models. The $\boldsymbol{\mu}$ and σ parameters for unmatched distributions remain the same. The parameters of the distribution that matches the new observation are updated as follows:

$$\begin{cases}\boldsymbol{\mu}_t=(1-\rho)\boldsymbol{\mu}_{t-1}+\rho\boldsymbol{X}_t,\\ \sigma^2=(1-\rho)\sigma_{t-1}^2+\rho(\boldsymbol{X}_t-\boldsymbol{\mu}_t)^{\mathrm{T}}(\boldsymbol{X}_t-\boldsymbol{\mu}_t),\end{cases} \quad (5)$$

where ρ is defined as

$$\rho=\alpha\eta(\boldsymbol{X}_t\mid\boldsymbol{\mu}_k,\sigma_k). \quad (6)$$

After the parameter updating process has ended, the Gaussians are ordered by the value of ω/σ. This value increases both as a distribution gains more evidence and as the variance decreases. This ordering of the model is effectively an ordered, open-ended list, where the most likely background distributions remain on top and the less probable transient background distributions gravitate towards the bottom and are eventually replaced by new distributions.

SUBSPACE FEATURE EXTRACTION METHODS

PCA (Martinez and Kak, 2001) is a method commonly used to represent an object. It selects projection directions of the datasets from maximal to minimal variances. When the projection matrix $\boldsymbol{w}^*$ is computed, if $\boldsymbol{x}$ is a new feature vector, then it is projected to $\boldsymbol{y}=\boldsymbol{w}^{*\mathrm{T}}(\boldsymbol{x}-\boldsymbol{m})$, where $\boldsymbol{m}$ is the average of samples. The vector $\boldsymbol{y}$ is used in place of $\boldsymbol{x}$ for representation.

To handle the complexity and variety of kinds of moving objects and background clutter, we tried to find a dimensionality reduction algorithm which can preserve the intrinsic features of different moving objects and classify them better. It seemed that linear discriminant analysis (LDA) and MFA may be more suitable. We will discuss these methods in this section and compare their performance in the experimental section.

Linear discriminant analysis

The aim of LDA (Martinez and Kak, 2001) is to find the directions that are most effective for discrimination by minimizing the ratio between the intraclass and interclass scatters. Given a number of training samples $\boldsymbol{x}_i^c \in \mathbb{R}^N$ in c known classes, where c is the class number and i is the sample ID in the cth class, the interclass scatters matrix $\boldsymbol{S}_\mathrm{b}$ and the intraclass scatters matrix $\boldsymbol{S}_\mathrm{w}$ are defined as

$$\begin{cases} \boldsymbol{S}_\mathrm{b} = \sum_{c=1}^{N_c} n_c (\overline{\boldsymbol{x}}^c - \overline{\boldsymbol{x}})(\overline{\boldsymbol{x}}^c - \overline{\boldsymbol{x}})^\mathrm{T}, \\ \boldsymbol{S}_\mathrm{w} = \sum_{i=1}^{N} (\boldsymbol{x}_i - \overline{\boldsymbol{x}}^{c_i})(\boldsymbol{x}_i - \overline{\boldsymbol{x}}^{c_i})^\mathrm{T}. \end{cases} \tag{7}$$

The projection is defined to minimize the ratio between the trace of $\boldsymbol{S}_\mathrm{w}$ and the trace of $\boldsymbol{S}_\mathrm{b}$:

$$\boldsymbol{w}^* = \arg\min_{\boldsymbol{w}} \frac{\boldsymbol{w}^\mathrm{T}\boldsymbol{S}_\mathrm{w}\boldsymbol{w}}{\boldsymbol{w}^\mathrm{T}\boldsymbol{S}_\mathrm{b}\boldsymbol{w}}. \tag{8}$$

The projection matrix $\boldsymbol{w}^*$ can be computed from the leading eigenvectors of $\boldsymbol{S}_\mathrm{w}^{-1}\boldsymbol{S}_\mathrm{b}$. LDA has the characteristic that the number of available projection directions is lower than the class number. So if we want to recognize only one object from the scene, we can obtain only one dimension projection. We showed in our experiment that the more that dimensions are reduced, the lower the recognition rate. It then becomes necessary to choose an appropriate dimension according to the recognition rate curves. Therefore, we introduced the MFA algorithm to obtain projection directions of moving objects. Many more projection directions are available from using MFA than from using LDA. They are determined by the selected number of shortest pairs of in-class and out-of-class sample pairs.

Marginal Fisher analysis

MFA was designed to characterize the intraclass compactness and the interclass separability by following the graph embedding formulation (Yan *et al.*, 2007). Here we briefly introduce its principle. As illustrated in Fig.2, the intraclass compactness graph $G=\{X, \boldsymbol{W}\}$ is constructed to show the intraclass point adjacency relationship. Each sample $\boldsymbol{x}_i \in X$ is connected to its k_1-nearest neighbors of the same class. If $\boldsymbol{x}_i$ is among the k_1-nearest neighbors of $\boldsymbol{x}_j$ in the same class, set the adjacency matrix $W_{ij}=W_{ji}=1$. This process is expressed as

$$\begin{aligned} S_c^m &= \sum_{i\in N_{k_1}^+(j)} \sum_{j\in N_{k_1}^+(i)} \left\|\boldsymbol{y}_i - \boldsymbol{y}_j\right\|^2 = \sum_{i\in N_{k_1}^+(j)} \sum_{j\in N_{k_1}^+(i)} \left\|\boldsymbol{w}^\mathrm{T}\boldsymbol{x}_i - \boldsymbol{w}^\mathrm{T}\boldsymbol{x}_j\right\|^2 \\ &= 2\boldsymbol{w}^\mathrm{T}\boldsymbol{X}(\boldsymbol{D}-\boldsymbol{W})\boldsymbol{X}^\mathrm{T}\boldsymbol{w}, \end{aligned} \tag{9}$$

$$W_{ij} = \begin{cases} 1, & \text{if } i \in N_{k1}^+(j) \text{ or } j \in N_{k1}^+(i), \\ 0, & \text{else,} \end{cases}$$

$$D_{ij} = \sum_{i\neq j} W_{ij},$$

where $\boldsymbol{y}_i$ is the dimension-reduced vector, and $N_{k_1}^+(i)$ indicates the index set of the k_1-nearest neighbors of the samples $\boldsymbol{x}_i$ in the same class.

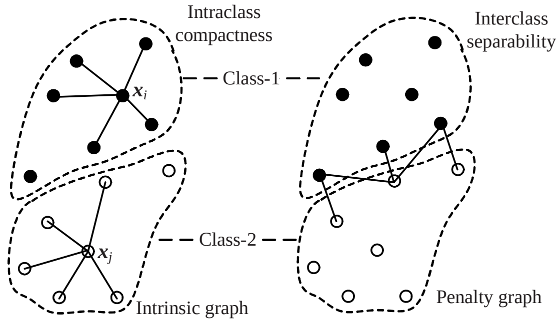

Fig.2 Adjacency relationships of the intrinsic and penalty graphs

A penalty graph $G^{\mathrm{p}}=\{X, \boldsymbol{W}^{\mathrm{p}}\}$, is defined to show that the interclass marginal point adjacency relationship and the marginal point pairs of different classes are connected. For each class c, if the pair (i, j) is among the k_2-shortest pairs among the set $\{(i, j), i\notin c_i\}$, set the similarity matrix $W_{ij}^{\mathrm{p}}=1$:

$$\begin{aligned} S_p^m &= \sum_{(i,j)\in P_{k_2}(c_i)} \sum_{(i,j)\in P_{k_2}(c_j)} \left\| \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i - \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_j \right\|^2 \\ &= 2\boldsymbol{w}^{\mathrm{T}}\boldsymbol{X}(\boldsymbol{D}^{\mathrm{p}}-\boldsymbol{W}^{\mathrm{p}})\boldsymbol{X}^{\mathrm{T}}\boldsymbol{w}, \end{aligned} \tag{10}$$

$$W_{ij}=\begin{cases} 1, \text{ if } (i,j)\in P_{k_2}(c_i) \text{ or } (i,j)\in P_{k_2}(c_j), \\ 0, \text{ else,} \end{cases}$$

where $P_{k_2}(c_i)$ is a set of data pairs that are the k_2-nearest pairs among the set $\{(i, j), i\in c_i, j\notin c_i\}$. Then the Marginal Fisher Criterion is

$$\boldsymbol{w}^* = \arg\min_{w} \frac{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{X}(\boldsymbol{D}-\boldsymbol{W})\boldsymbol{X}^{\mathrm{T}}\boldsymbol{w}}{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{X}(\boldsymbol{D}^{\mathrm{p}}-\boldsymbol{W}^{\mathrm{p}})\boldsymbol{X}^{\mathrm{T}}\boldsymbol{w}}. \tag{11}$$

Compared with other discriminant methods, MFA is designed to characterize interclass margin and intraclass compactness. There are no assumptions about the data distribution for each class. The intraclass compactness is characterized by the sum of the distances between each data point and its k_1-nearest neighbors of the same class. The available projection directions are determined by k_2 which is the selected number of the shortest pairs of in-class and out-of-class sample pairs. So it is not limited by the class number, and the interclass margin defined in MFA can better characterize the separability of different classes than the interclass variance in LDA.

EXPERIMENTS AND DISCUSSION

Moving objects database establishment

As a general moving objects recognition method, any kind of moving object can be extracted and built into the moving objects database. Automobiles and pedestrians are the most typical and important objects within the application. Here we divided the segmented objects into three categories: automobiles, pedestrians, and others. The database should comprise enough samples of automobiles, pedestrians and other moving objects to enable it to describe the kinds of objects within the subspace feature. In most research work where the data sample is small, the subspace distribution is not representative and the results are not convincing. Our automobile database comprises all sorts of motor vehicles including cars, buses, trucks, etc. The pedestrian database includes walking people viewed from a range of visual angles. The remaining objects belong to another database and include bicycles, animals, partially occluded objects, clutter, etc.

We have built our sample database from video recordings collected from a roadside environment. Once the background distribution is established, the foreground pixels that do not match the distribution are extracted. Connected components are segmented and unified to 50×50 (pixels) square. The unifying process does not distort the connected components, but patches them into a square. The moving objects are put into the middle of the square and it is then resized to 50×50. Thus a moving object silhouette image is extracted and unified, as illustrated in Fig.3.

Fig.3 Samples of automobiles, pedestrians and the other objects

We separated the sample database into a training set and a testing set. Subspace features were obtained with the training set using dimensionality reduction algorithms. The testing set was used to verify the recognition rate of our method. The numbers of samples in the training and testing sets are listed in Table 1.

Table 1 Number of moving objects in the database

Moving objects	Number of training sets	Number of testing sets
Pedestrians	4276	872
Automobiles	4568	857
Others	9031	1314

Comparison of different subspace features

For convenience of illustration, we have reduced the moving object database to 2D. Fig.4 shows that pedestrians and automobiles are more concentrated compared with the other objects. If we want to distinguish only one object from another object, it is very easily achieved using PCA. Fig.5 shows the 2D distributions of pedestrians and automobiles, which are separated completely. We want to be able to distinguish a certain object from the background clutter. Therefore, we call our method 'moving objects recognition', but not 'moving objects classification'.

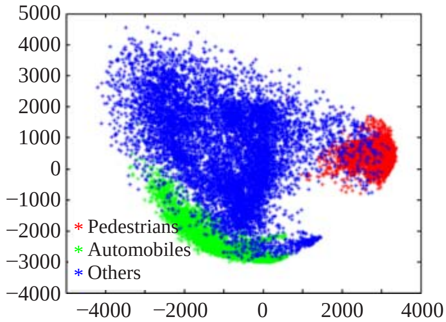

Fig.4 Dimensionality reduction results by PCA

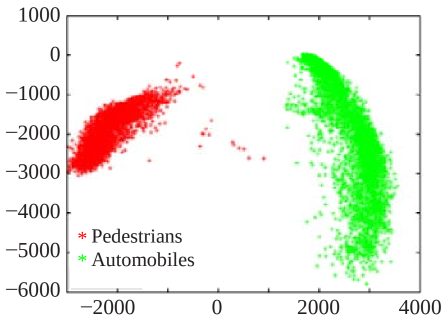

Fig.5 Dimensionality reduction results for two sets by PCA

Fig.6a shows the dimensionality reduction results of the training set using LDA, and Fig.6b shows the result using MFA with certain k_1 and k_2. We find that different classes are more clearly separated by LDA and MFA than by PCA. In the next subsection, we compare their classification ability with their recognition rate curves.

Comparison of recognition performances of different subspace features

The recognition performances based on our test sets were computed to verify the effectiveness of our moving objects recognition algorithm, and the

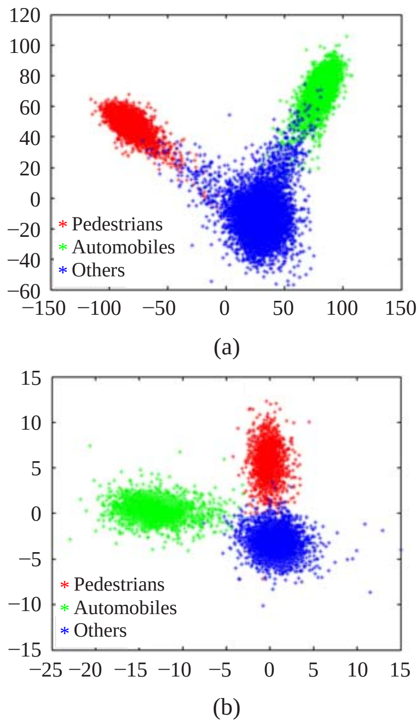

Fig.6 Dimensionality reduction results by LDA (a) and MFA (b)

recognition performances of different dimensionality reduction algorithms were compared. KNN was used as the classifier to compute the recognition rate, and here K was set to 10. Figs.7a~7c are the recognition rate curves of PCA and MFA on pedestrians, automobiles and other objects, respectively. The x-axis represents the datasets' dimensions retained by the dimensionality reduction algorithm. The y-axis represents the different recognition rates at each dimension.

Because the maximum available projection directions of LDA is 2D in our experiment, we compared the recognition rates of PCA, LDA and MFA in 2D (Fig.8).

We conclude from the recognition rate results that our method is feasible and effective for the moving objects recognition problem. By characterizing the intraclass compactness graph and the interclass penalty graph of the training set, MFA can obtain better performance in this classification problem compared with the classical linear dimensionality reduction methods of PCA and LDA. This is a significant improvement in a real application.

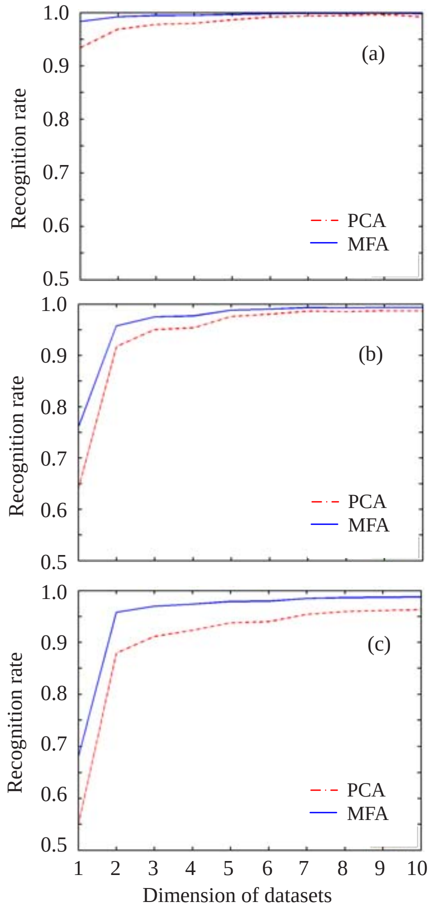

Fig.7 Recognition rates of pedestrians (a), automobiles (b) and others (c)

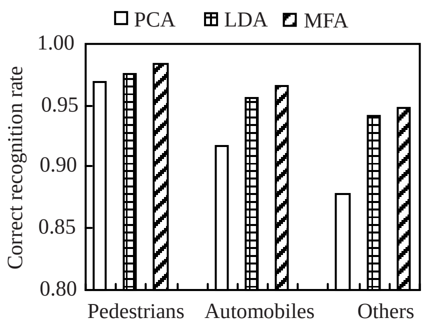

Fig.8 Comparison of recognition rates of PCA, LDA and MFA

DISCUSSION

Our system framework of moving objects recognition has no rigorous demands for background subtraction algorithms, while some algorithms need a continuous silhouette image of the moving objects (Lipton *et al*., 1998; Wang and Tan, 2003). Therefore, our system is more applicable. However, the algorithm used for the training database establishment should be the same as the algorithm used in the recognition system. We find that the recognition rates reduce too much if a different background subtraction algorithm is used.

From the recognition rate curves, we can choose a suitable dimension value which represents subspace features of different classes and is sufficient to get a good recognition result. In this experiment, we can find three dimensions that are the most suitable values.

k_1 and k_2 represent the nearest neighbors of $\boldsymbol{x}_j$ in the intraclass compactness graph and the interclass penalty graph which are the two key parameters in the MFA method. It is noticeable that a good recognition result is based on suitable choices of k_1 and k_2. Fig.9 shows the dimensionality reduction result from using MFA within another group of k_1 and k_2. It shows that the different classes cannot be classified.

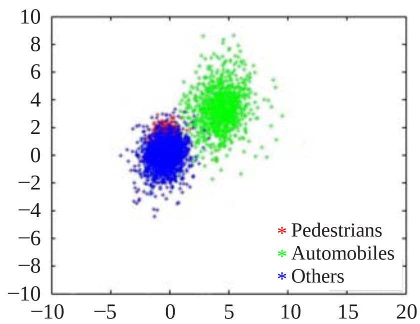

Fig.9 Dimensionality reduction results using MFA with another group of k_1, k_2

We have experimented with different automobile types for recognition when an object has been classified as an automobile. The automobile set is classified into cars, jeeps, trunks, buses and sports cars. The training and recognition processes are just the same. The recognition results were also reassuring and gave us great confidence in our general moving object recognition method.

In real applications, when one object overlaps another, or one is partially occluded by other objects, its silhouette image will be connected with the others or damaged. In our system, it may then be classified in the 'others' class. Thus, it will be lost in the tracking process. This is a common problem in detection and recognition algorithms. In our system, a mean shift tracking algorithm is used to solve this problem (Comaniciu *et al*., 2000). If we have detected an automobile in the current frame and want to track it continuously, the automobile's color distribution is modeled as the target model. Then mean shift iteration is used to find the most probable target position in the next frame. This is a robust tracking algorithm in real applications. Figs.10 and 11 show a process of recognition result via background subtraction only. The automobile connects with the human in Fig.11. Fig.12 shows the result when the mean shift tracking algorithm is used.

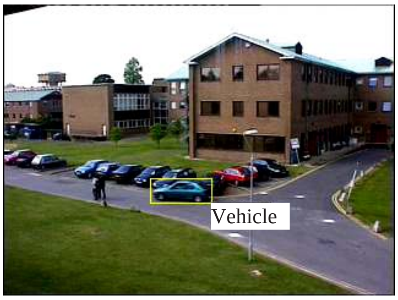

Fig.10 Recognition result with background only

Fig.11 Recognition result with background only (Automobile connected with human)

Fig.12 Recognition result with mean shift tracking algorithm

CONCLUSION

In this study, we have presented a general method for a moving objects recognition system. Based on large numbers of real samples, we can obtain precise object recognition results with a subspace feature method. This method was more general and robust compared with classical object feature extraction methods which aim at particular objects. It also is not limited by whether the object is rigid or non-rigid. Using this framework, recognition performance can be significantly improved by using the MFA algorithm, as illustrated in our experiment. Based on this recognition system, video surveillance such as abnormal behavior detection or access control to special areas, can be further developed.

References

Collins, R.T., Lipton, A.J., Kanade, T., Fujiyoshi, H., Duggins, D., Tsin, Y., Tolliver, D., Enomoto, N., Hasegawa, O., Burt, P., *et al*., 2000. A system for video surveillance and monitoring. Tech. Rep., CMU-RI-TR-00-12, Carnegie Mellon University, Pittsburgh, PA.

Comaniciu, D., Ramesh, V., Meer, P., 2000. Real-time Tracking of Non-rigid Objects Using Meanshift. IEEE Conf. on Computer Vision and Pattern Recognition, p.142-149.

Cutler, R., Davis, L.S., 2000. Robust real-time periodic motion detection, analysis, and applications. *IEEE Trans. Pattern Anal. Mach. Intell.*, **22**(8):781-796. [doi:10.1109/34.868681]

Duque, D., Santos, H., Cortez, P., 2007. Prediction of Abnormal Behaviors for Intelligent Video Surveillance Systems. Proc. IEEE Symp. on Computational Intelligence and Data Mining, p.362-367. [doi:10.1109/CIDM.2007.368897]

El Baf, F., Bouwmans, T., Vachon, B., 2007. Comparison of Background Subtraction Methods for a Multimedia Application. 6th EURASIP Conf. on Speech and Image Processing, Multimedia Communications and Services, p.385-388. [doi:10.1109/IWSSIP.2007.4381122]

Ghobadi, S.E., Hartmann, K., 2006. Detection and Classification of Moving Objects—Stereo or Time-of-flight Images. Int. Conf. on Computational Intelligence and Security, p.11-16. [doi:10.1109/ICCIAS.2006.294082]

Haritaoglu, I., Harwood, D., Davis, L.S., 2000. W4: Real-time surveillance of people and their activities. *IEEE Trans. Pattern Anal. Mach. Intell.*, **22**(8):809-830. [doi:10.1109/34.868683]

Hu, W.M., Tan, T.N., 2004. A survey on visual surveillance of object motion and behaviors. *IEEE Trans. Syst., Man, Cybern.—Part C: Appl. Rev.*, **34**(3):334-352. [doi:10.1109/TSMCC.2004.829274]

Javed, O., Ali, S., 2005. Online Detection and Classification of Moving Objects Using Progressively Improving Detectors. Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, p.696-701.

Lipton, A.J., Fujiyoshi, H., Patil, R.S., 1998. Moving Target

Classification and Tracking from Real-time Video. 4th IEEE Workshop on Applications of Computer Vision, p.8-14. [doi:10.1109/ACV.1998.732851]

Liu, X.D., Su, G.D., 2000. A New Network-based Intelligent Surveillance System. 5th Int. Conf. on Signal Processing Proc., p.1187-1192.

Martinez, A.M., Kak, A.C., 2001. PCA versus LDA. *IEEE Trans. Pattern Anal. Mach. Intell.*, **23**(2):228-233. [doi:10.1109/34.908974]

Munder, S., Gavrila, D.M., 2006. An experimental study on pedestrian classification. *IEEE Trans. Pattern Anal. Mach. Intell.*, **28**(11):1-6. [doi:10.1109/TPAMI.2006.217]

Rivlin, E., Rudzsky, M., 2002. A Real-time System for Classification of Moving Objects. 16th Int. Conf. on Pattern Recognition, p.688-691.

Stauffer, C., 1999. Automatic Hierarchical Classification Using Time-base Co-occurrences. Proc. IEEE Conf. on Computer Vision and Pattern Recognition, p.335-339.

Stauffer, C., Grimson, W.E.L, 1999. Adaptive Background Mixture Models for Real-time Tracking. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, p.246-252.

Wang, L., Tan, T.N., 2003. Silhouette analysis-based gait recognition for human identification. *IEEE Trans. Pattern Anal. Mach. Intell.*, **25**(12):1505-1518. [doi:10.1109/TPAMI.2003.1251144]

Yan, S.C., Xu, D., Zhang, B.Y., 2007. Graph embedding and extensions: a general framework for dimensionality reduction. *IEEE Trans. Pattern Anal. Mach. Intell.*, **29**(1):40-51. [doi:10.1109/TPAMI.2007.250598]