# Pedestrian Tracking by using MPEG-based Video Signal Processing

Mikio SASAKI† Hideaki NANBA†

†Research Laboratories, DENSO CORPORATION 500-1 Minamiyama, Komenoki-cho, Nisshin-shi, Aichi, 470-0111 Japan

E-mail: †{ msasaki@rlab.denso.co.jp, HIDEAKI_NANBA@denso.co.jp}

**Abstract** In this paper, we propose a method for pedestrian detection and tracking which is based on MPEG video processing. Firstly AC power of frame difference for each block of pixels is calculated through DCT coefficients and is evaluated as a confidence measure of the pedestrian. Secondly, similarity of the motion between neighboring MPEG macroblocks is calculated and evaluated as a classification measure for moving objects. Finally these two measures are integrated and judged by the specified rule base to output recognition results. This first stage method has independently attained the maximum recognition rate of more than 90% per second. As the second stage algorithm for the moving camera, a multivariate time series analysis using the PCA (Principal Component Analysis) has been tried. Also future works on adaptation and behavior modeling are investigated.

**Keyword** Pedestrian tracking, MPEG, DCT, Motion, PCA, Confidence, Behavior, Time series analysis, Video processing

## 1. INTRODUCTION

In the last decade, many of vehicle related companies have devoted themselves to developing highly assured safety systems. However, most of them are ongoing technologies and few systems have been industrialized. Especially this is true on on-board traffic recognition systems. What makes it difficult is mainly ascribed to the gap between the level of perfection that many car makers demand and the level of realistic technologies regarding pedestrian detection and behavior understanding. In this aspect, we proposed a probe car based pedestrian prediction system [1]. However we have found that the recognition accuracy of pedestrian by using conventional methods does not reach our target level even if it is improved by our original scene recognition techniques. This is mainly because static image based pedestrian detection is insufficient for human recognition especially in the driving situations. Therefore we focused on the motion pictures recognition and also on the combination with wireless communication technologies. One of the proper solutions is MPEG-based video recognition which already started in the area of broadcasting and multimedia. In this paper, we propose an application of the technologies to the pedestrian detection from compressed video data transmitted through on-board wireless communication equipments.

## 2. RELATED WORK

In the field of the pedestrian recognition, there are many approaches and several kinds of sensor fusion systems [2], [3], [4]. In view of sensing, there exist various devices such as laser radar, millimeter-wave radar, ultrasonic device, video camera, IR camera, stereo camera, etc. Here, we especially focus on monocular vision.

In the area of image processing, there exists a lot of pedestrian recognition methods such as the combination of SVM (Support Vector Machine) and Ada Boost [5], neural network based method, model based method, template matching and appearance based methods, etc. Those are classified into snapshot based techniques and usually do not utilize motion or temporal characteristics of image sequence that are quite important to detecting human in the ambiguous or noisy signal environments.

On the other hand, recently various human tracking methods have been developed [6], [7], [8]. Some of them are triggered by serious security requirements and additionally, those sometimes refer further to behavior understanding, non-stationary abnormality detection [8], and human trajectory analysis [9]. Moreover, crowds recognition is becoming important and fundamental research has been started [10].

From another viewpoint, there are also step-by-step development works that are aimed at reduction of misdetection [1], solution to occlusion problems [7], and acquisition of robustness [6]. We already introduced our integrated method of confidence vector based scene recognition (CVSCR) and pattern based pedestrian detection (PPD) such as neural network. This is because PPD independently cannot discriminate pseudo human and also has some various defects regarding occlusions, shape changes, and common sense judgments.

According to those backgrounds, we have focused on video sequence recognition towards pedestrian tracking.

## 3. TARGETED SYSTEM

In view of system, the targeted application determines which method is the most appropriate one for us. Generally speaking, as is well known, there are three levels of safety such as crash safety, active safety, and danger anticipation. A captured scene would be classified into one of these categories on the basis of TTC (Time To Collision). Currently many vehicle related companies are making efforts to develop crash safety oriented systems. However, the more the TTC decreases, the more likely the

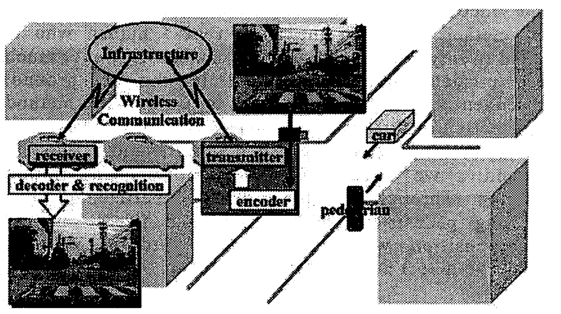


Fig. 1. Our Targeted System..

crash becomes inevitable. So we have focused on crash anticipation phase in which TTC would be larger than about five seconds. On the other hand, in such a phase, on-board camera would not be able to capture pedestrians sufficiently because of image size and various occlusion problems. Therefore, wireless communication infrastructures are useful and we have considered the target system shown by Fig. 1. This is aimed at detecting the approach of a pedestrian who is occluded and/or situated within dead angle, which enables warning the driver in advance. In this case it becomes possible for the vendors to separate productive responsibilities on recognition performance from sensors and communication infrastructures. We believe this point will pave the way for the realistic industrialization.

This is why we can focus on image processing based pedestrian tracking. Regarding image recognition, we have already tried snapshot based pedestrian detection [1], however we have found that method cannot reach the performance we need although more than eighty percent of misdetections can be reduced by using our original scene recognition techniques. Moreover, the still picture based techniques in principle cannot capture the dynamic activities of human and also cannot follow the abrupt change of motion. This is the basic reason why we have introduced motion pictures based recognition methods like MPEG-based pedestrian detection.

From other industrial aspects, MPEG has special interoperability over multimedia, internet, and broadcasting industries according to the position of the worldwide ISO standards. And in the last two decades, no doubt it has already been backed up by huge amounts of sophisticated technologies including software, hardware, and systems. These derive mainly from the feature that high quality of video versus compression rate is achieved. This nature strongly suits for wireless communication and mobile equipments.

## 4. FIRST STAGE ALGORITHM

Generally in MPEG based video coding, frame difference is largely converted to two major video coding categories. First one is that DCT (Discrete Cosine Transform) coefficients which are well known as the most appropriate approximation of KLT (Karhunen-Loeve Transform) when assuming video signal statistics. And second one is that motion vectors that represent two dimensional interframe positional difference on block of pixels. So roughly speaking, we can identify the moving objects and scene change by judging these two features.

### 4.1. Preliminary Experiments on Motion

At first, it has been assumed that any kinds of object motions would be detected by using excellent motion detection algorithm. However this assumption has been turned down through the preliminary experiments on the famous KLT (Kanade-Lucas-Tomasi) method [11] for base band video signals. The results showed it is possible for KLT method to detect the motion of a human who is around twenty meters distant from camera. However much more misdetections were found in the background. Consequently, it would be difficult to use the KLT method for detection of distant human because no prospective and filtering methods were established at that stage.

Secondly, we tried MPEG-1 based human motion detection. Through our preliminary experiments when capturing pedestrians from more than twenty meters distant position, we found that it is hard for MPEG-1 encoder to detect human motion precisely.

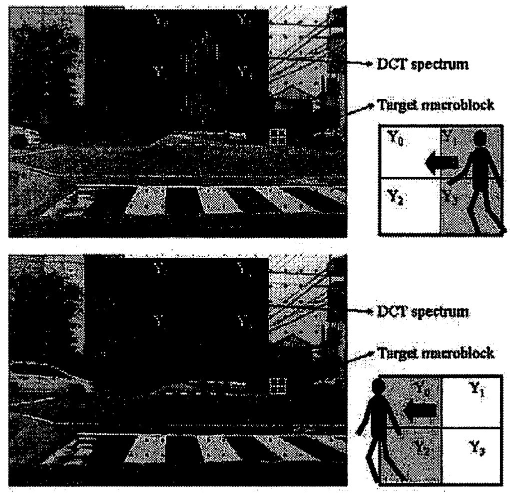


Fig. 2. Pedestrian detection by using DCT spectrum.

### 4.2. DCT Coefficients

Therefore, we have decided to focus on temporary characteristics of DCT coefficients because it is considered that no other signals could express the change of objects if no motion has been detected.

Firstly, we have set observation window the size and position of which correctly match to one MPEG-1 macro block that includes four blocks of Y component pixels of 8×8. Upper picture in Fig, 2 shows the scene in which one person is entering into this window from right. Four graphics in the upper middle area of the picture is showing two dimensional DCT coefficients for each block of pixels in the observation window. For simplicity we call the window as BEF (Block Edge Filter). As can be easily seen, the right half of the BEF is showing remarkable levels of DCT coefficients which express human motion. This feature does not depend on the background textures because MPEG-based DCT coefficients are showing DCT spectrum of frame difference for a block of pixels. Lower picture in Fig. 2 shows that the human is about to exit the BEF and the corresponding DCT spectrums are prominent in the left hand side of the BEF. Through the characteristics of Fig. 2, we easily find that power evaluation of DCT coefficients in each BEF would be a practical way to find a moving human.

A basic evaluation function for AC power within each block of pixels is defined as follows:

$$L_{AC}(l_{blk}) = \sum_{m=0}^{63}\sum_{n=0}^{63}\left|DCT(l_{blk},m,n)\right| - \left|DCT(l_{blk},0,0)\right| \qquad (1)$$

Where, $DCT(l_{blk}, m, n)$ means two dimensional DCT coefficient in which $l_{blk}$ denotes block number in MPEG macroblock (hereafter, we call MBK). And $m$ and $n$ mean horizontal coordinate, vertical coordinate, respectively.

Then we can define the sets of threshold values for judging the sum of absolute values of DCT coefficients according to the picture coding types that are defined in the ISO/MPEG standard. Basically, there are at least three types of picture coding types in MPEG standard such as I (intra), P (predictive), B (interpolational). In P or B frame, there exists MBKs that are coded as interframe difference or motion compensated difference over 16×16 pixels.

Within the MBK, the BEF can detect the moving objects which propagate prominent components in 256 dimensional spatio-temporal waves of DCT coefficients.

**4.3. Motion Evaluation**

As can be easily seen in Fig. 3, a moving vehicle generates MBK based motion vectors in MPEG-1 format, when it is less than twenty meters distant and horizontally traversing the image plane. In this figure, blue arrow means forward motion vector $\boldsymbol{v}_f$ which is extracted from the past frame and the current frame. Also red arrow means backward motion vector $\boldsymbol{v}_b$ which is extracted from the future frame and the current frame according to the MPEG-1 standard. For simplicity, we will consider the synthesis of these two motion vectors as the total motion vector $\boldsymbol{v}$ for the corresponding MBK as defined below.

$$v(l_{mbk}) = v_b(l_{mbk}) - v_f(l_{mbk}) \tag{2}$$

Where, $l_{mbk}$ means number of MBK in a picture. According to this definition, we can easily find that each MBK included in a vehicle area tends to have the motion vectors of similar magnitude and similar direction. These tendencies do not occur on pedestrian area unless the pedestrian is very close to the camera. These empirical results can be formalized as the rules for judgment.

*1) Magnitude of Motion:* Generally, a normal pedestrian walks much slower than vehicle and number of pixels of interframe change is fairly small. A distant pedestrian tends to generate almost no motion vectors or a few pixels of magnitude at most. Therefore we can discriminate a vehicle from a human by using the following inequality.

$$|v| \geq M_{car} \tag{3}$$

Where, $M_{car}$ denotes the threshold for judge and $|.|$ denotes the Euclid distance or absolute distance. We have used $M_{car}$ = 4 for the case of a vehicle that runs horizontally at less than 30 km/h and at around 20 meters distant location.

The similarity of motion in magnitude is calculated by

$$SAMV = 1 - \frac{\left\| v_1 | - | v_2 \right\|}{\left(|v_1| + |v_2|\right)} \tag{4}$$

This is normalized to take a value between 0 and 1.

*2) Direction of Motion:* Direction of motion can be restricted if the shape and the road plane are known in advance. The similarity of motion in direction can be defined as follows:

$$SDMV = \frac{(1+\cos\theta)}{2} = \frac{1}{2}\left(1 + \frac{v_1 v_2^T}{|v_1||v_2|}\right) \tag{5}$$

Where, $v_i$ (i=1,2) is a two dimensional motion vector and $\theta$ denotes the angle between $v_1$ and $v_2$. *SDMV* is normalized to take a value between 0 and 1.

*3) Total Similarity of Motion:* For two adjacent MBKs in *i*-th MBK line, we can define the total similarity of motion SMV as follows:

$$SMV(i,j) = \frac{v(i,j)v(i,j+1)^T}{|v(i,j)||v(i,j+1)|} \tag{6}$$

Where, *v(i,j)* means a two dimensional row vector for a motion of *j*-th MBK in *i*-th MBK line in a picture.

**4.4. Rule Based Decision**

When performing final judgment of moving objects, there exist a lot of cases in which simple numerical judgments described above do not sufficiently work. From our various empirical results on real road scenes, we believe there should be many rules that are specifically applied to road scene and on-board camera. The rules which seem to be particular effective and relatively easy to implement are as follows:

- Vehicle tends to have almost no textures on surface of body. However boundary parts between body and tires often make AC components and it leads to the drastic change of DCT coefficients in BEF and will cause misdetection of human.
- Human tends to have many shadow parts and also have complicated textures.
- Vehicle motion tends to be continuous and to have less up and down. On the other hand, human has some specific and periodical motion.
- Generally, vehicle moves much faster than human.
- BEF of vehicle often makes complete occlusion.
- BEF of human often shows background textures.
- BEF of human depends on cloths that affect DCT spectrum and silhouette.
- Human often makes non-Gaussian motions that include abrupt stop, direction change, etc.

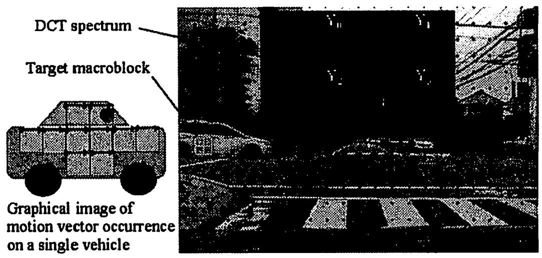


Fig.3. Vehicle detection by using motion vectors.

**4.5. Towards Real Time System**

Fig.4 shows the architecture of MPEG-based recognition system which suggests the real time implementation. In our method, the MPEG video encoder LSI which is widely spread in many kinds of electronics equipments can be considered as the most reliable real time processor for image features extraction. If a slight modification for the data interface is possible, we will not need to have parsing operation of MPEG streams in recognition process. This is because we can extract MPEG features directly from MPEG encoder processor. This expectation will reduce the realization cost drastically.

## 5. EXPERIMENTS AND RESULTS

**5.1. Conventional Method**

As we have shown in ITST2007 [1], neural network based detection is used as a conventional method. But it will not show good performance for tracking because of two reasons. First reason is that it only performs frame

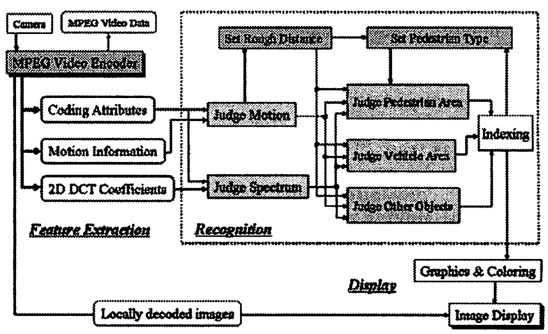


Fig. 4. Architecture of MPEG-based video recognition.

Fig. 5. Experimental results.

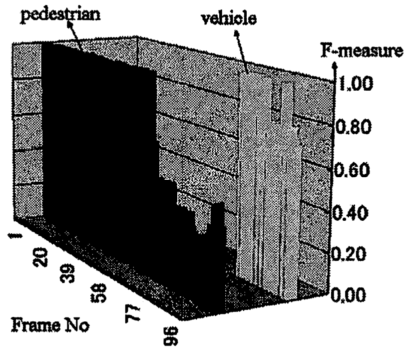


Fig. 6. Temporal characteristics of F-measures.

independent pattern detection. Second reason is that it is not good at recognition of dynamic behaviors of pedestrians according to frame by frame learning of a specific representative shape.

**5.2. MPEG-based Method**

Fig.5 shows the experimental results of our proposed MPEG-based method. Every block of pixels which is judged as a part of pedestrians is colored by red. Similarly every block of pixels which is judged as a part of moving vehicles is colored by yellow. Fig.6 shows the temporal characteristics of F-measure [1] corresponding to the experimental results in Fig.5. During the first two seconds (60 frames), pedestrian tracking rate (red) is more than 90%/sec and maximum average is 94%/sec. In the latter part after the two seconds, the performance is degraded. This is mainly because a white car approaching the pedestrian from the left side is detected as shown by yellow blocks and a several number of misrecognized red blocks (pedestrian) occur on the edges of the vehicle. Simultaneously, vehicle tracking rate (yellow) surges up after the first two seconds and shows more than 90%/sec on average.

Regarding misrecognized blocks, there are at least four types. First type of blocks which is found on the edges of vehicles would be deleted by incorporating more sophisticated decision logic which utilizes shape information. Second type of misdetection which is found on sky or building etc. would be deleted by using scene recognition technique such as CVSCR [1]. Third type which shows blinking in a few frames would be a noise. Other types such as included in far distant moving vehicles needs more close analysis. The same kinds of problems will arise when we challenge the scene captured by a moving on-board camera.

**5.3. Evaluation**

Similar trial for the same sequence by using the conventional pedestrian detection method (neural network based pattern recognition for each independent frame) showed less than 29 %/sec for tracking rate in F-measure. Also from the aspects of subjective evaluation of video output, the conventional method cannot perform sufficient tracking of moving objects. Therefore it is concluded that our proposed MPEG-based video recognition highly surpassed the conventional method of frame independent pattern recognition for the crossing case.

**5.4. Study Towards Second Stage Algorithm**

We define the second stage as application to moving camera. In order to realize it, the basic equation (1) on the basis of AC power should be modified to approach closer to human image characteristics. There are at least two strategies for that. The first one is to incorporate statistical time series analysis. Second one is to extract effective components from 64 DCT coefficients. Therefore, as the preliminary experiments for integration, we tried multivariate time series analysis using principal component analysis (hereafter, PCA).

Firstly, let the feature vector be composed from $M_{walk}$ consecutive predictive frames in MPEG format. Here, the $n$-th ($n=1,\dots,N$) feature vector is defined as follows:

$$\boldsymbol{x}_n = \left(x_{n1},\cdots,x_{nP}\right)^T \tag{7}$$

Where, $P$ means the number of dimensions for the feature vector. As we tentatively set $M_{walk}$ to five, the duration for the time series is 0.5 second in our experiment. Then $N$ becomes 64×5=320.

Corresponding measurement matrix is denoted as below.

$$X = \left(\boldsymbol{x}_1 \cdots \boldsymbol{x}_N\right)^T \tag{8}$$

This matrix should be defined on the data extracted from pedestrian scene in order to perform effective learning of human image characteristics.

Then, the covariance matrix $S$ is defined as below.

$$S = \frac{1}{N}X^T X \tag{9}$$

The $p$-th ($p=1,\dots,P$) eigenvector of $S$ is denoted by

$$\boldsymbol{w}_p = \left(w_{p1},\cdots,w_{pP}\right)^T \tag{10}$$

Then the principal component scores are calculated by using next formula.

$$\boldsymbol{f}_p = X\boldsymbol{w}_p \tag{11}$$

This is easily extended to the normalized matrix form as follows:

$$\begin{aligned} F &= \left(\boldsymbol{f}_1 \cdots \boldsymbol{f}_P\right) \\ &= XW diag(1/\sqrt{\lambda_1}\cdots 1/\sqrt{\lambda_P}) = XW\Delta_S^{1/2} \end{aligned} \tag{12}$$

Where, $W = \left(\boldsymbol{w}_1,\cdots,\boldsymbol{w}_P\right)$ and $\lambda_p$ denotes $p$-th eigenvalue of $S$, namely, $\Delta_S \equiv diag\left(\lambda_1,\cdots,\lambda_P\right)$.

Fig.7 shows the actual flow of recognition process on the basis of these calculations above. Remark that the parametric eigenspace is utilized to visualize the effectiveness of the largest three principal components during a pedestrian scene period. According to that, we tentatively defined the new evaluation function as the absolute distance between origin and the each point on the trajectory in the eigenspace. Therefore we can calculate the confidence measure using this distance. The results in Fig.7 are showing fairly good recognition performances for two different scenes.

Additionally, we have tried another method which is called spatio-temporal image as shown by Fig.8. The fascinating feature of this method is that it becomes possible to extract gait information easily by using relatively simple image processing techniques. The crossing waves observed on the image show pedestrian's walking behavior. The corresponding waving shapes were also observed on the differential signal based spatio-temporal image even in the case of moving camera, according to our experiments.

## 6. FUTURE WORKS

Fig. 9 shows the adaptation mechanism towards more sophisticated algorithm that can recognize the scene even if on-board camera is moving. There are two strategies to adapt dynamically the parameters for judgments. First one is statistical scene analysis [1] which enables the estimation of environments and driving status. Second strategy is to incorporate pedestrian modeling in which basic behaviors are described by a state transition diagram as shown by Fig.10.

### 6.1. Behavior Model of Pedestrian

Behavior understanding plays an important role in prediction of the danger which driver is to come across. From the aspects of driving safety, pedestrian behaviors are classified into three classes such as normal, dangerous, abnormal. Those classes are closely related each other as can be seen in Fig.10. By using these relations, adaptation mechanism in Fig.9 will work well to keep high recognition rate. Also it would be possible to identify more accurately the pedestrian profile which includes class of act, status of attitude, class of behavior, age and gender, clothes, hair style, baggage, etc. even if signal source is ambiguously sensed information.

### 6.2. Detection of Abnormal Behavior

Following cases are considered as abnormal.

*1) Jumping Out:* Generally, pedestrians are on the pavement, tend to be upright position and rarely lean. And a pedestrian autonomously appears, moves, and disappears through the door or stairs except the case of occlusion. Provided that the first appearance is observable, there would be five cases for "jumping out" as follows:

a) To jump out straightforward after walking or running through open pavement that has no occlusions.

b) To jump out straightforward after walking or running through rather occluded pavement because of trees or vehicles or other pedestrians.

c) To jump out straightforward after standing still for a while on the open pavement that has no occlusions.

d) To jump out straightforward after standing *still* for a while on rather occluded pavement because of trees or vehicles or other pedestrians.

e) To jump out straightforward after standing still for a while in the middle of the vehicle road.

In these cases, a) and e) would be detectable by using our method described above. Detection possibility of b) and c) highly depends on the degree of occlusions. However it would not be impossible through the use of hidden state tracking. For the case of d), behavior transition model in Fig.10 is fairly effective and from the aspects of human dynamics, "leaning forward" should be observed previously. Therefore if such an attitude change can be detected as some of low frequency vertical components of DCT coefficients, the following behavior "jumping out" would become predictable and detectable.

*2) Standing Still:* In the pavement, "standing still" is not dangerous but, in the middle of the road, it often turns to be dangerous behavior. In view of DCT coefficients, "standing still" will cause the drastic decrease of AC power of equation (1). Therefore the interframe difference in AC power defined as below is to be evaluated.

$$D(n) = L_{AC}(n) - L_{AC}(n-1) \tag{13}$$

Where, $n$ denotes the discrete time number for a frame. This can be performed by a logical equation which includes two inequalities as follows:

$$A = \{D(n) < 0\} AND \left\{ \left[ \sum_{k=0}^{K} L_{AC}(n+k) \right] \leq E_{th\_stop} \right\} \tag{14}$$

Where, '*AND*' means logical product and $E_{th\_stop}$ means an empirically determined threshold value. $K$ means an appropriate length of observation period in discrete time. Then if the logical value $A$ is true, it can be judged as "standing still" through the continuous observation.

### 6.3. Detection of Omen

*1) Large Motion:* Empirically, it is natural to consider that large motion around vehicle road is an omen of abnormal behaviors, particularly when it is peculiar as

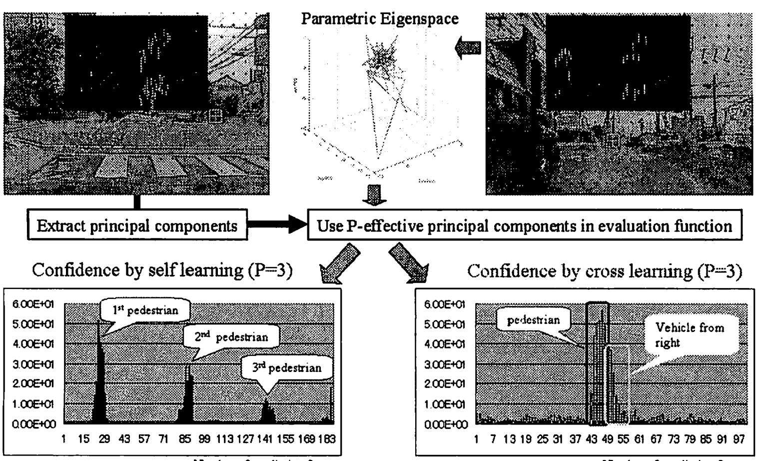


Fig. 7. Approach from multivariate time series analysis.

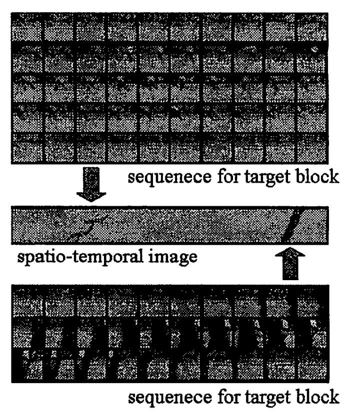


Fig. 8. Spatio-temporal image.

compared with others. Following cases are considered:

a) Large mobility (running, walking fast, etc.)

In this case, motion vectors are detectable even if the target pedestrian is around 20 meters distant. However, its magnitude rarely becomes larger than that of vehicle. This case sometimes generates intra-coded block rather than motion compensated difference block.

b) Large movement of arms and legs

There also would be much difference in BEF and it generates higher AC power than that of ordinary walking.

c) Brandishing flags or poster or sign board

This case would generate more interframe differences in BEF that leads to high level of DCT coefficients.

*2) Vertical Motion:* If prominent DCT coefficients occur periodically in vertically adjacent BEFs, there would be possibilities that pedestrian repeats up and down behaviors like hop or jump. This class of vertical behaviors is to be judged as symptoms of abnormal behaviors and corresponding confidence values can be set higher than before.

## 7. CONCLUSION

We proposed MPEG-based pedestrian recognition method. When the on-board camera stops at the crossing, it attained maximum tracking rate of more than 90 % for each of pedestrian and vehicle. This performance completely outperforms the conventional snapshot based method. Our method is so much suitable for wireless communication that we can construct the safety information system which imposes no responsibilities on sensing devices, transmitters, and infrastructures.

We are now under development of the second stage algorithm which can attain high performance even if on-board camera is moving fast by using statistical video analysis and behavior modeling.


## ACKNOWLEDGMENT

The authors would like to thank Dr. Tadahiro Kitahashi who eagerly joined our discussions and gave us precious comments especially regarding PCA related topics.


## REFERENCES


[1] M. Sasaki, K. Muto, K. Sasaki, "Estimation of Driving Safety from On-Board Captured Scene by using Probe Image Database", *ITST2007*, June 2007.

[2] M., Marc-Michael; et al., "SAVE-U: First Experiences with a Pre-Crash System for Enhancing Pedestrian Safety". http://www.save-u.org/download/PDF/ITS_Hannover_200506.pdf.

[3] D. M. Gavrila, S. Munder, "Multi-cue Pedestrian Detection and Tracking from a Moving Vehicle", *International Journal of Computer Vision* 73(1), 41-59,2007.

[4] J. Tilp, et al., "Pedestrian Protection based on Combined Sensor Systems", http://www-nrd.nhtsa.dot.gov/pdf/nrd-01/esv/esv19/05-0156-O.pdf

[5] P. Viola, M. J. Jones, "Robust Real-time Object Detection", *Cambridge Research Laboratory Technical Report Series CRL2001/01, Feb. 2001.*

[6] J.-M. Odobez, D. Gatica-Perez, and S. O. Ba, "Embedding Motion in Model-Based Stochastic Tracking", *IEEE Trans. Image Processing*, vol.15, no.11, pp.3514-3530, Nov. 2006.

[7] Bo Wu, Ram Nevatia, "Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet based Part Detectors", *International Journal of Computer Vision* Volume 75, Issue 2 (November 2007), 247 – 266.

[8] Xiang, Tao; Gong, Shaogang; "Beyond Tracking: Modeling Activity and Understanding Behaviour.", *International Journal of Computer Vision*, Vol. 67, No. 1, April 2006 , pp. 21-51(31).

[9] G. Antonini, "A discrete choice modeling framework for pedestrian walking behavior with application to human tracking in video sequences", *Thèse EPFL*, no 3382 (2005).

[10] M. Lyell, R. Flo, M. Mejia-Tellez, "Simulation of Pedestrian Agent Crowds, with Crisis", http://necsi.org/events/iccs6/viewpaper.php?id=367 .

[11] Jianbo Shi and Carlo Tomasi. Good features to track. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 593–600, 1994.


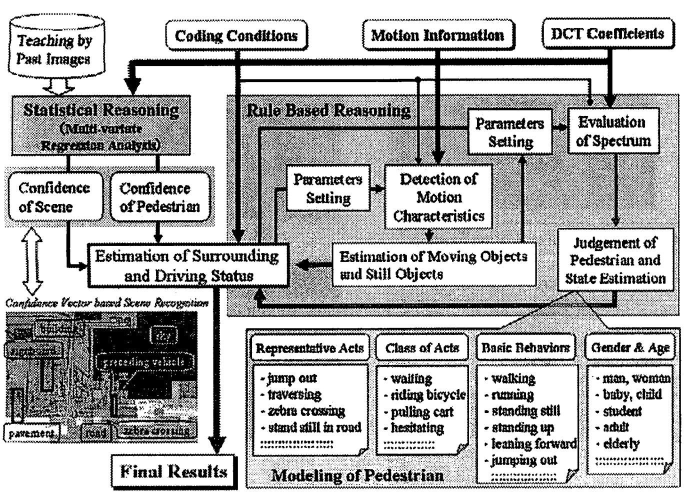


Fig. 9. Adaptation mechanism.

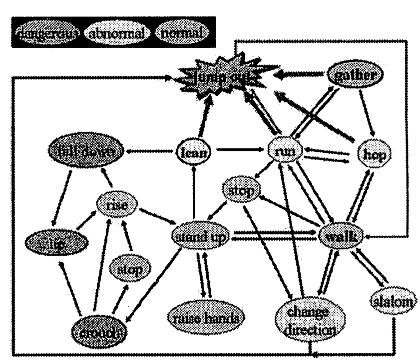


Fig. 10. Modeling of pedestrian