Moving Object Tracking Using Co-occurrence Features of Objects

Seongdong Kim
Dept. information & communication network at Kaywon School of Art and Design
(sdkim@kaywon.ac.kr)

Seongah Chin
Division of Multimedia, Sungkyul University
(esolideo@sungkyul.edu)

Moonwon Choo
Division of Multimedia, Sungkyul University
(mchoo@sungkyul.edu)

In this paper, we propose an object tracking system which can be convinced of moving area shaped on objects through color sequential images, decided moving directions of foot messengers or vehicles of image sequences. In static camera, we suggests a new evaluating method extracting co-occurrence matrix with feature vectors of RGB after analyzing and blocking difference images, which is accessed to field of camera view for motion. They are energy, entropy, contrast, maximum probability, inverse difference moment, and correlation of RGB color vectors. we describe how to analyze and compute corresponding relations of objects between adjacent frames. In the clustering, we apply an algorithm of FCM(fuzzy c means) to analyze matching and clustering problems of adjacent frames of the featured vectors, energy and entropy, gotten from previous phase. In the matching phase, we also propose a method to know correspondence relation that can track motion each objects by clustering with similar area, compute object centers and cluster around them in case of same objects based on membership function of motion area of adjacent frames.

Key words: Object tracking, Motion tracking, Object motion detection, Co-occurrence

1. Introduction

Moving object tracking in video sequences is considerably important as related hardware/ software technology becomes more effieient and the needs for applications where the activity of objects should be analyzed and monitored are increasing [1], [3], [6], [7], [10], [12], [13], [15]. In such applications lots of information can be obtained from trajectories that give the spatio-temporal coordinates of each objects in the environment. Information that can be obtained from such trajectories includes a dynamic count of the number of object within the monitored area, time spent by objects in an area and traffic flow patterns in an environment. The tracking of moving object is challenging in any conditions, since image formations in video stream is very sensitive to changes of conditions of environment such as illumination, moving speed and directions of objects, the number and sizes of objects, and background. Therefore the scope of researches are usually confined to specific application domains and the processes of capturing video streams are also carefully controlled.

In this paper, general moving objects are considered and color information is utilized as main source for extracting features for tracking since gray-level images may lose much of information available in color space such as combined and synthesized features derived from separate color channels.

We suggest a system for obtaining such spatio-temporal tracks of objects in video sequences. Camera in static position produces video sequences which are analyzed in real time to obtain trajectories. In each frame of video stream, segmentation technique such as differencing gray-level intensities embedded in inter-frame images could well work in real time and yield regions of interest for blocking quickly. An important step towards the track of objects is the definition of a proper set of features, which could reliably identify the corresponding objects in adjacent frames. Most of the motion tracking researches focus on the features generated from the gray-level images which generally derived from the RGB space. Color image can be assumed to contain richer information for image processing than its gray-level image. Also separate color channel could be applied to different problem domains. In this paper, color space is exploited to extract the reliable features for tracking moving objects.

Color co-occurrence matrices generated by special position operator to conjecture the color patterns embedded in target image blocks is exploited in this paper. The objects in video sequence usually has the additional unique color patterns which could not be detected in gray-level ones. The descriptors of color co-occurrence matrices, such as energy, entropy, contrast, maximum probability, inverse different moment, and correlation, are analyzed to determine the availability as feature vector in order to identify adjacent block images as the identical objects and find out the trajectories of objects. Fuzzy C- Means algorithm is used to determine the correspondence of adjacent two image blocks by computing the distance from the central point of fuzzy clusters formed in feature spaces.

2. Related Research

Jakub and Sarma[11] developed a system for realtime tracking of people in video sequences. They use a model-based approach to object tracking, identifying feature points like local curvature extrema in each video frame. Their system has an advantage of handling occlusion problems, but disadvantage of unreliable extraction of extrema of curvature from object contours. Other approaches suggest motion tracking by deriving velocity vectors from point-to-point correspondence relations. Relaxation and optical flow are very attractive methodologies to detect the trajectories of objects. Those researches are based on the analysis of velocity vectors of each pixel or group of pixels between two neighboring frames. This approach requires heavy computation for calculating optical flow vectors. Another method infers the moving information by computing the difference images and edge features for complementary information to estimate plausible moving tracks [3],[5]. This method may be very sensitive to illumination and noise imposed on video stream. The other method adopts the model-based approach, which has disadvantage of extracting the previously trained objects only.

3. Object tracking system

A flow chart showing the main steps is given in <Figure 1>.

3.1 Block detection

Block detection is a primarily important stage in object tracking applications. Many blocking methods assume that the lighting in the scene considered would be constant. The accuracy of these methods decreases significantly when they are applied to real scenes. A multiple-level blocking method based on threshold differences and a morphological focus of attention able to reduce the effects of noise and of changes in lighting is suggested.

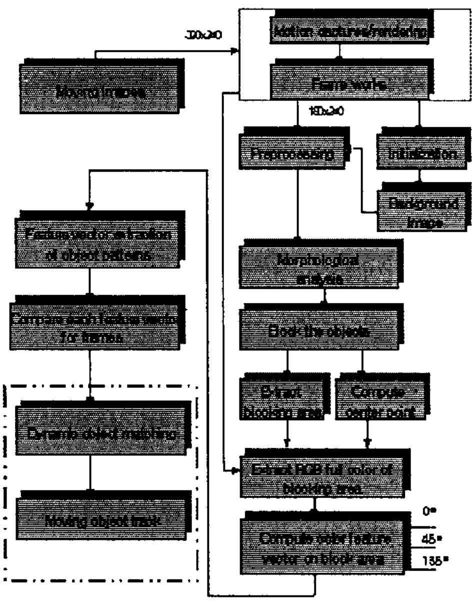
This module receives a pair of gray-level image frames, $I_{k-1}(x,y)$ and $I_k(x,y)$, acquired at successive time instants t_{k-1} and t_k, respectively. Then a list of minimum bounding rectangle- shaped blocks of image areas where significant changes (related to possible moving objects) is produced.

The proposed module consists of three steps.

1. Computing the difference $D_k(x, y)$, between the two input images $I_{k-1}(x,y)$ and $I_k(x,y)$,

   $$D_k(x, y) = \mid I_k(x, y) - I_{k-1}(x, y) \mid .$$

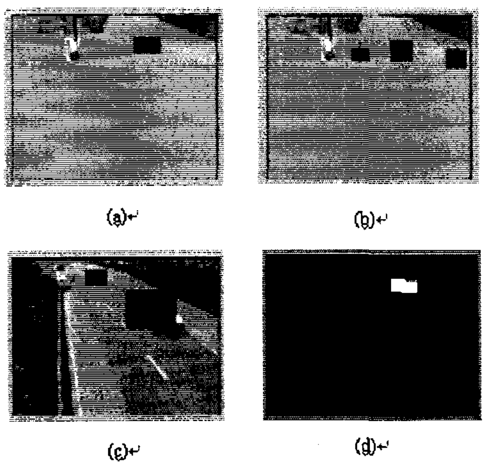
2. Establishing whether each point $(x, y) \in D_k(x, y)$, is a background point or a moving-object point, then generating the binary block image B. The function used for point labeling is a spatial hysteresis function designed to improve the algorithm robustness [6]. If the state is background and $D_k(x, y) > \theta_{in}$ the state of the point is switched to object. If the state is object and $D_k(x, y) > \theta_{out}$ the state of the



<Figure 1> Main steps in the moving object tracking system

point is switched to background. The selection of threshold values, θ_{in} and θ_{in} are strictly dependent on the geometry of the image frame considered such as dimensions of a moving object and distance between moving objects and camera, etc.

Noise filtering and searching for the minimum bounding rectangular shaped blocks are performed by means of simple morphological operation and searching extremal points.



<Figure 2> (a) a real image representing a road scene with overlapped walking persons and its detected block (b) separated waking persons (c) moving vehicles (d) dilation image of (a).

<Figure 2> shows a real image representing a road scene with multiple people and vehicles, and detected blocks.

3.2 Moving Object Counting

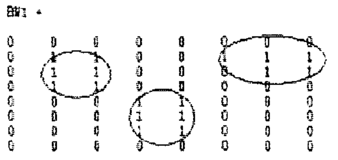
Up to now, we notice that moving objects can be easily detected by blocking. We, however, are not able to count the number of moving objects for a scene, which make it possible to compare the features specified in the next section between objects from two consecutive frames. By usage of connected component into the blocked image after morphological operation, we figure out how many objects appear in a frame. The operation is working upon detected blocking image $B(x, y)$ as shown in <Figure 2> (d) conveying only moving objects. <Figure 3> (b) demonstrates that three moving objects can be clearly detected via a connected component algorithm given blocking binary image (a),

3.3 Color feature extraction

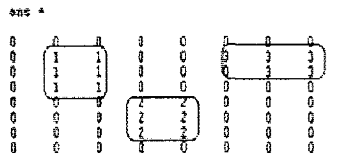
The presence of object detected with its bounding rectangular block is associated with the RGB values imposed on corresponding pixels. Among all statistical methods, the most popular one, which is based on the estimation of the second-order statistics of the spatial arrangement of the gray level values, is the gray-level co-occurrence matrices[9]. A co-occurrence matrix $P_{\theta,d}(a, b)$ for a displacement d, θ direction, and gray level values a and b is a matrix whose elements correspond to the relative frequency of occurrence of pairs of gray level values of pixels separated by a certain distance in a given direction. Formally, the elements of $M \times N$ gray level co-occurrence matrix $P_{\theta,d}(a, b)$ for a displacement d, $\theta=0$ direction, and gray level values a and b, is defined as[13].

$$P_{0^{\circ},d}(a,b) = \left| \left\{ \begin{array}{l} [(k,l),(m,n)] \in D : k - m = 0, |l - n| = d, \\ f(k,l) = a, f(m,n) = b \end{array} \right\} \right|$$

where | | refers to set cardinality and $D = (M \times N)(M \times N)$.



(a)



(b)

<Figure 3> (a) a binary image of three blocking objects (b) extracted three connected components

In this paper, the displacement d is set to 1 and the only 0°, 45°, 135° directions are considered since other directions may show symmetrical properties. Also the color co-occurrence matrices corresponding to each RGB plane is computed. From this matrix, several textural features, each of which represents certain image properties such as coarseness, contrast, homogeneity and texture complexity, could be derived. Those features that are used in this work are:

(1) Energy/angular second moment :

$$\sum_{a,b} P^2_{\phi,d}(a,b)$$

Energy gives a measure of the homogeneity of an image. Hence it is a suitable measure for detection of disorders in textures. For homogeneous textures the energy value turns out to be small compared to non-homogeneous ones.

(2) Entropy:

$$\sum_{a,b} P^2_{\phi,d}(a,b)\log_2 P_{\phi,d}(a,b)$$

Entropy gives a measure of complexity of the image. Complex textures tend to have higher entropy.

(3) Maximum probability

$$\max_{a,b} P_{\phi,d}(a,b)$$

(4) Contrast

$$\sum_{a,b} |a-b|^k P^\lambda_{\phi,d}(a,b)$$

Contrast feature is a measure of the image contrast or the amount of local variations present in an image.

(5) Inverse difference moment:

$$\sum_{a,b,a \neq b} \frac{P^\lambda_{\phi,d}(a,b)}{|a-b|^k}$$

(6) Correlation:

$$\frac{\sum_{a,b}[(a,b)P_{\phi,d}(a,b)] - \mu_x\mu_y}{\sigma_x\sigma_y},$$

where, $\mu_x\mu_y$ are means and $\sigma_x\sigma_y$ are standard deviations.

Correlation is a measure of image linearity. Linear directional structures in direction ϕ result in large correlation values in this direction.

3.4. Object tracking

The goal of this module is to determine the 3-D positions and the motion parameters of objects recognized by the blocking module, at every time instant. Each detected object block may be assumed to contain only one moving object. But this constraint can be alleviated, since the textural features could be jointly utilized to differentiate multiple objects possibly overlapped.

To establish the correspondence relations of blocks between sequential frames, the color features embedded in co-occurrence matrices obtained from each block are used as input to the tracking module.

The FCM (fuzzy c-means algorithm) has been used for the clustering and classification of feature vectors [7]. It attempts to cluster measurement vectors by searching for local minima of the generalized within group sum of squared errors functions. The FCM is given by

$$J_m(u,v)=\sum_{k=1}^{n}\sum_{i=1}^{c}u_{ik}{}^{m}\left|x_k-v_i\right|_A^2$$

where c is the number of clusters, n is the number of vectors, x_k is a kth measurement vector, v_i is the i th centroid vector, $m\in[1,\infty]$ is the fuzzy coefficient, $|.|_A$ is an inner product norm, $|Q|^2_A=Q^TAQ$, and A is a $d\times d$ positive definite matrix where d is the dimension of the pattern vectors. Here, $u=\{u_{ik}\}$, where u_{ik} is a membership function with input values. Our FCM starts with selecting the number of clusters c, determined by the number of connected components in a scene. The method to obtain the number of object is described in the section 3.2. In this paper,

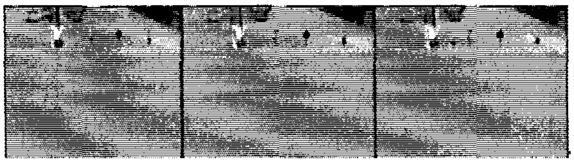
$$x_k=(X_r,X_g,X_b)_\theta,$$
$$v_i=(V_r,V_g,V_b)_\theta,\quad \theta=0^\circ,45^\circ,135^\circ$$

where X_r, X_g, X_b are the feature vectors of RGB components respectively and v_i is the centroids of each RGB components.

To demonstrate moving object trajectory, we adapt the central object point calculated by central moment. After investigating correspondence of various objects between consecutive frames, each object draws its trajectory using its central object point.

4. Experimental results

The color features suggested here are extensively analyzed to determine the plausibility of candidacy as input to tracking module. Several video clips are generated using static digital video camera according to the kinds of objects and their speed variations under natural illumination conditions. Also the variety of different values of parameters of co-occurrence matrix such as distance and angle are examined.



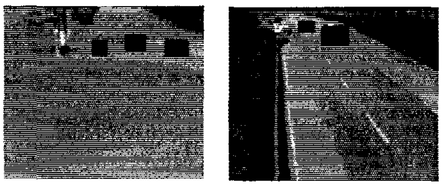
(a) (b) (c)

<Figure 4> Three sequential image frames of walking persons

<Table 1> shows the one of results after detecting blocks of sequential image frames from <Figure 3>.

4.1 Effects of color features on the correspondence relations

The color features should allow ignorance of a larger set of disturbing factors and more discriminative power comparatively. To find out the eligibility as discriminative feature vectors, the color features of certain objects are computed and compared numerically with same object class and different object class as well. <Figure 5> shows two major images to be tested. <Figure 6> shows the part of results of color features of two different objects. As a result, we may conclude that the color features except energy and entropy have less discriminative power, even though the object classes are different.



(a) (b)

<Figure 5> The two video sequences to be tested (a) walking persons (b) cars.

We have discovered importantly fundamental fact that features of energy and entropy fall into quite acceptable threshold bound between two timely consecutive same objects whereas the features stray for different classes as shown in <Figure 6> and <Figure 7>.

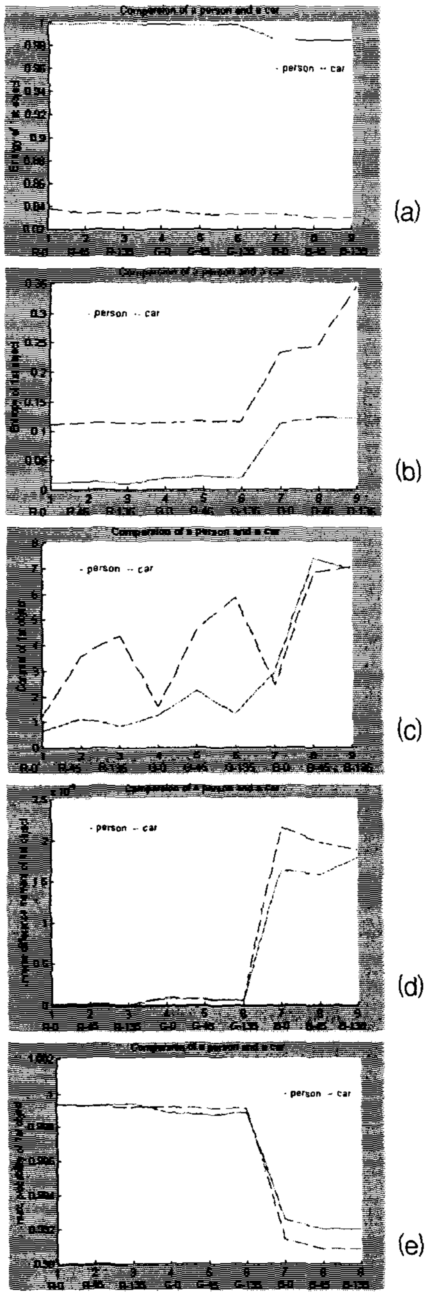
In the long run, we have selected only two features (energy and entropy) for dynamic matching stage.

4.2 A Motion tracking by clustering

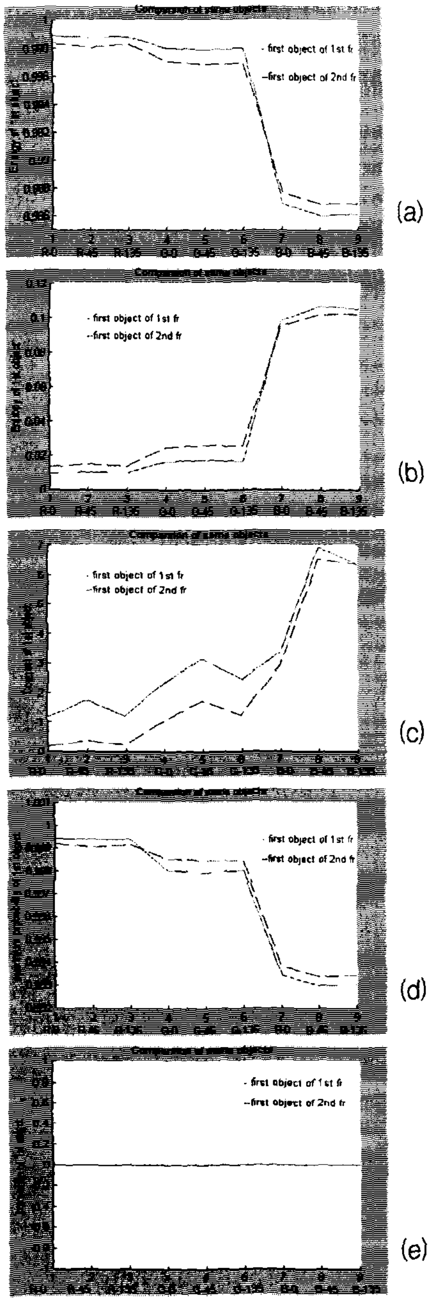
The energy and entropy features are applied to detect the tracks of moving objects. FCM gives the reasonable clustering mechanism to identify the correspondence relations. <Figure 8> shows the clustering of the energy and entropy color feature of three moving walkers. <Figure 9> and <Figure 10> demonstrate whole procedures of the proposed algorithm sequentially. <Figure 9> (i) and <Figure 10> (e) show the resultant moving tracks of three persons.

5. Conclusions and future research

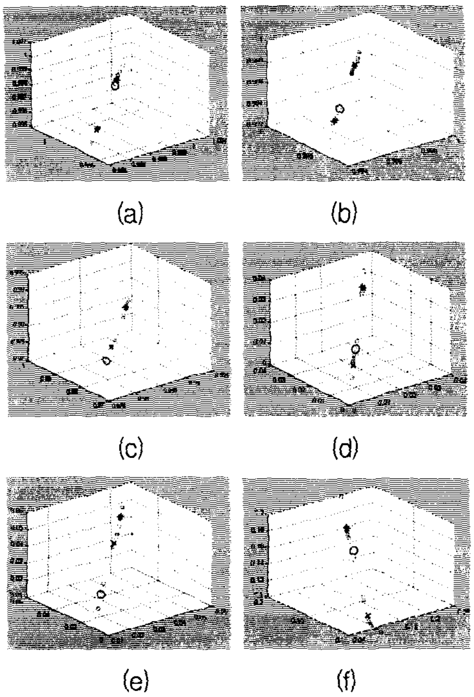
A system for tracking objects using color features is presented. The system outputs tracks that give spatio-temporal coordinates of objects as they move within the field of view of a camera. We try to solve the occluded objects problem, which may be ignored in many researches intentionally and the utilization of color features numerically. The problematic issues imposed on



<Figure 6> The color feature vectors of two different object classes. (a) energy (b) entropy (c) contrast (d) inverse DM (e) Max. Prob.

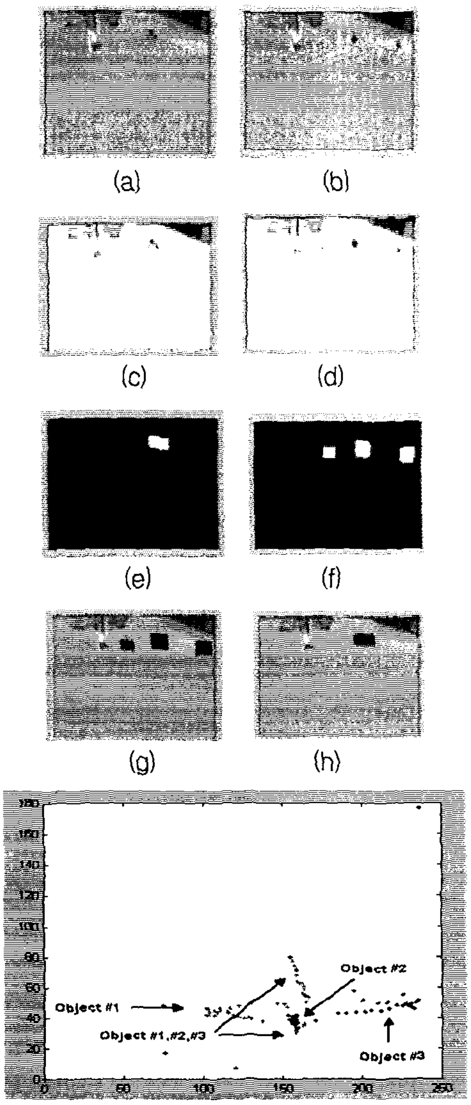


<Figure 7> The color feature vectors of same object classes. (a) energy (b) entropy (c) contrast (d) Max. Prob. (e) Correlation



<Figure 8> Clusters of energy feature of R(a), G(b), B(c) planes (the number of objects = 3) and clusters of entropy feature of R(d), G(e), B(f) planes (the number of objects = 3)

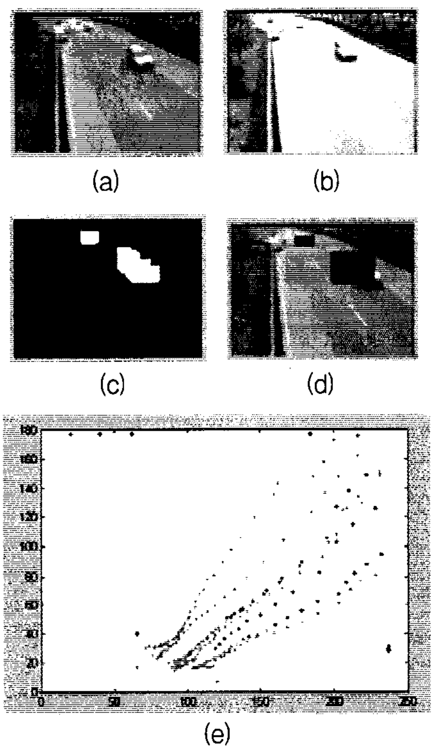
this system are time complexity which are caused many color feature vectors associated with FCM and ignorance of color invariance which may be prerequisite to any attempt to utilize the color features. Future research needs to address these kinds of issues and tracking objects across moving cameras.



<Figure 9> (a) (b) Input images, (c) (d) equalization of input images, (e) (f) dilation images, (g) (h) block of input images (i) trajectory of moving pedestrian

<Table 1> The results calculated separately from R planes for three directional texture patterns where d=1 and θ=0, 45, 90.

frame	1^{st} object (1^{st} frame)			1^{st} object (2^{nd} frame)			1^{st} object (3^{rd} frame)		
Color	R			R			R		
Degree	0	45	135	0	45	135	0	45	135
Energy	0.9986	0.9985	0.9987	0.9988	0.9987	0.9988	0.9983	0.9981	0.9983
Entropy	0.0114	0.0126	0.0107	0.0097	0.0105	0.0102	0.0135	0.0154	0.0137
Max pro.	0.9993	0.9992	0.9993	0.9994	0.9993	0.9994	0.9981	0.9990	0.9991
Contrast	0.6531	1.1366	0.8279	1.1427	1.7462	1.1962	0.1857	0.3651	0.2390
Inverse DMoment	0.00001	0.00001	0.000007	0.000003	0.000002	0.000004	0.00003	0.00004	0.00004
Corr.	0	0	0	0	0	0	0	0	0



<Figure 10> (a) Input video images, (b) equalization of input images, (c) dilation images, (d) block of input images (e) trajectory of moving pedestrian

References

[1] Beymer, D and Konolige, K, "Real-Time Tracking of Multiple People using Stereo," Proc. IEEE Frame Rate Workshop, 1999

[2] Chanda, B., Chaudhuri, B. and Dutta, Majumder "On image enhancement and thread selection using the greylevel co-occurrence matrix", Pattern recognition Lett. Vol.3, no.4, pp. 243-251, 1985

[3] Cucchira, Rita, Piccardi, Massimo, and Mello, Paola "Image analysis and rule based reasoning for a traffic monitoring system" IEEE, Intelligent transportation system, pp119-pp130 VOL.1, No.2, June 2000

[4] Culter, Ross and Davis, Larry S. "Robust Real-Time Periodic Motion Detection, and analysis and Application", IEEE Pattern analysis and machine Intelligence Vol22 No 8, pp781-795 August 2000

[5] Darrell, Trevor and Covel, Michele "Correspondence with cumulative similiarity transforms" IEEE pattern analysis and machine intelligence, pp 222-227 Vol.23 No2 February

2001

[6] Foresti, Gian Luca, et. al, "Vehicle Recognition and Tracking from Road Image Sequences," IEEE Trans. On Vehicular Tech., vol. 48, NO. 1, Jan. 1999.

[7] Haritaoglu, Ismail, David Harwood, Larry S. Davis, "W^4:Real time Surveillance of people and their activities", IEEE Trans. on pattern analysis and machine intelligence, Vol. 22, No. 8, Aug. 2000

[8] Haralick Robert M. and Shapario, Linda G., "Computer and Robot vision Vol I", Addision-wesley, USA pp 318-321, 1992

[9] Latif, A., et. al, An efficient method for texture defect detection: sub-band domain co-occurrence

[10] Rigoll, G. Eickeler S. and MullerPerson, S., Person Tracking in Real-World Scenarios Using Statistical Methods, IEEE2000 0-7695-0580-5/00 pp342-347

[11] Segen, Jakub and Pingali, Sarmai, "A Camera-Based System for Tracking People in Real Time," IEEE Proceedings of ICPR '96, 1996.

[12] Chin, Seongah "Image text segmentation", Ph.D dissertation 1999, Stevens Institute of Technology

[13] Sonka, Milan, Hiavac, Vaclav, and Boyle, Rogger, "Image processing Analysis and machine vision" International Thomson 1999, 2nd edition

[14] 김계영 "동영상 분석에 의한 실시간 이동물체 추적시스템", 박사학위논문, 숭실대 1997

[15] 이규원 "Bayes 결정과 특징기반움직임 플로우를 이용한 이동물체의 검출 및 추적", 박사학위논문, 연세대 1995

[16] 이장명 외 3인 "특징기반 다중 물체추적 시스템에 관한 연구", 대한전자공학회 논문지 Vol 36-S pp95-pp101 1999

[17] 이재구 외 3인 "무인감시장치 구현을 위한 단일 이동물체 추적 알고리즘", 대한전자공학회논문지 제32권 B편 11호 pp41-pp51 1995

요약

이동 물체의 상호 발생 특징정보를 이용한 동영상에서의 이동물체 추적

김 성 동**
진 성 아*
주 문 원*

본 논문에서는 연속적으로 입력되는 칼라영상에서 물체의 이동에 의하여 형성된 동작영역을 확인하고, 영상의 시컨스(sequence)를 대상으로 움직이는 물체의 형태인 보행자 혹은 자동차들의 이동방향을 추적하는 시스템을 제안하였다. 카메라가 고정되어 있고 물체가 이동하는 상황에서 카메라시계에 진입하는 물체를 포착하여, 포착된 물체의 영역을 차 영상 분석을 통해 이진화하여 추출하고, 추출된 영역을 co-occurrence matrix의 RGB full 칼라의 특징벡터를 추출하는 것을 제시하였다. 추출되어지는 칼라 특징벡터를 분석하여 인접 프레임간의 이동물체 영역끼리의 대응관계를 조사함으로서, 이동물체를 추적한다.

군집화(clustering) 단계에서는 이전 단계에서 추출한 특징 벡터들 가운데 에너지, 엔트로피만을 가지고 인접 프레임간의 군집화를 조사하기 위하여 이동물체 영역들 간의 퍼지동적물체 정합 알고리즘을 적용시켰다. 인접 프레임간의 움직임 영역의 물체들에 대하여 멤버 쉽 함수를 근거로 중심 값을 계산하면, 동일 물체일 경우 중심 값 부근에서 군집이 형성되며, 이를 바탕으로 이동물체를 추출할 수 있는 방안을 제안하였다.

* 성결대학교 멀티미디어학부
** 계원조형예술대학 정보통신과