Outdoor Human Tracking using Spatio-Temporal Information

Jun-ichi YAMAGUCHI and Nobue KOBAYASHI
Engineering Institution
Sogo Keibi Hosho Co., Ltd.
Sokei Riverside Bldg. 2-14 Ishijima, Koutou-ku, Tokyo, 135-0014
E-mail:yamaguchi@sok.co.jp
JAPAN

Commission V, Working Group IC V/III

KEY WORDS: Human Finding, Tracking, Motion Vector, Hough Transform, Image Processing

ABSTRACT

In outdoor human finding using image processing, there is a problem of detection error in case of a background change caused by a branches and leaves trembling in the wind, a light reflection on the surface of the water and so on. Therefore, the automatic human finding is generally restricted within the adaptation to the scene of less disturbance. This paper describes an outdoor human tracking method which is uninfluenced by disturbance, using a spatio-temporal information of the image. In this method, the motion vector is detected by computing an image changing region. The position of the motion vector is shown by two parameters (the angle θ and the distance ρ), and the shifting quantity is voted to the θ-ρ space. The voting data obtained from passer-by is concentrated on a local region in the θ-ρ space. On the other hand, the voting data obtained from disturbance shows a tendency to be distributed at random. The existence of the human is detected by estimating the vote data which continues at the same position in the θ-ρ space. To track the human, the motion data which contributes to the continuous vote data is extracted. Tracking the coordinates of center-gravity of the extracted motion data, a locus of the human is detected. The results of the experiment, which was performed to verify the effectiveness of the proposed method, are demonstrated.

1. INTRODUCTION

The human tracking is useful for understanding a pattern of behavior, counting the number of passer-by, security and so on. Many methods on the human finding have proposed up to the present, and recently some of them are put to practical use. However, those methods are generally restricted within the adaptation to the scene of less disturbance. Particularly, the methods for outdoor are restricted, because of the influence of the disturbance which is a light changing, a branches and leaves trembling in the wind, a light reflection on the surface of the water and so on [1][2]. Therefore, it is required that the automatic human finding method has both the elimination of the influence of the disturbance and the elimination of the undetected error.

The authors propose an outdoor human tracking method which is uninfluenced by the disturbance, using a spatio-temporal information of the image. This method finds the passer-by, using the straight line detection ability and noise elimination ability of Hough Transform [3]. In the method, the motion vector is detected by computing the image changing region. The position of the motion vector is shown by two parameters (the angle θ and the distance ρ), and the shifting quantity is voted to the θ-ρ space. If a peak of the voted data continues at the same position in the θ-ρ space, the existence of the human is detected. To track the human, the motion data which contributes to the peak is extracted. Tracking the coordinates of center-gravity of the extracted motion data, a locus of the human is detected.

This paper describes the method for tracking the outdoor human, using the spatio-temporal information of the image. The results of the experiment, which was performed to verify the effectiveness of the proposed method, are demonstrated.

2. HUMAN TRACKING ALGORITHM

2.1 DETECTION OF EXISTENCE

The changing region of the image is extracted by comparing two images which are obtained at a constant time interval. The motion vector is detected by computing the correlation about the image changing region in a search area, using a matching method [4][5]. In human detection, there are some cases that it is difficult to detect the change of the image because of a speed, a moving course, a distance from the camera and so on. In such case, the detectable passer-by and the detectable area are restricted. And so, in order to eliminate such restriction as much as possible, the changing region is extracted by three steps as shown in fig.1. The motion data in the step, which obtains more motion data than other two steps, is used for following processing. In detection of the motion data

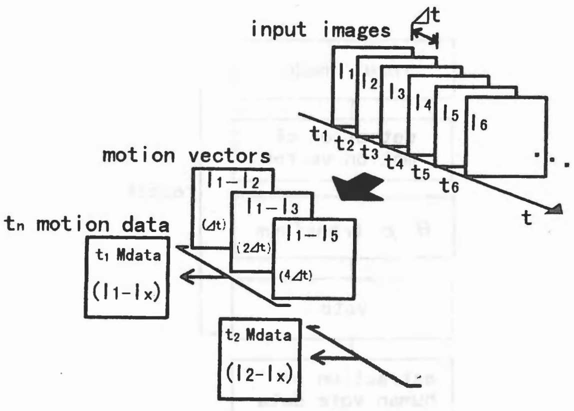


Fig.1 Detection of the motion data

Fig.2 An example of the motion vector

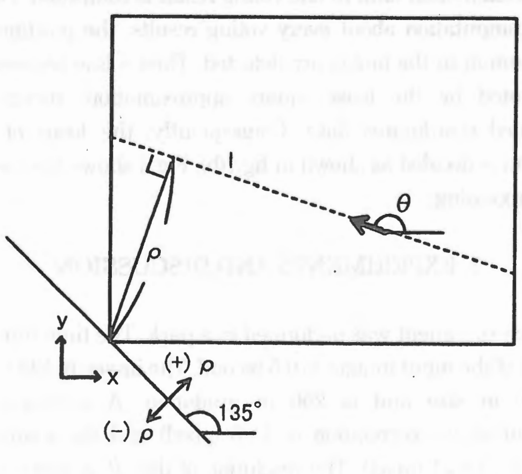


Fig.3 θ - ρ transform on motion vector

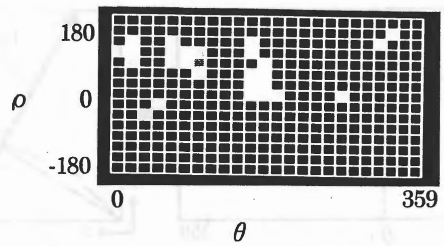


Fig.4 Result of vote (using the motion data in fig.2)

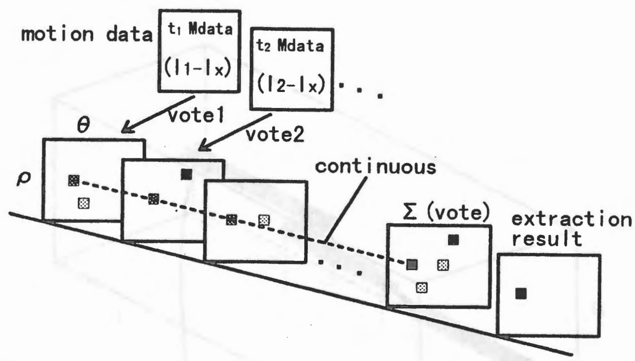


Fig.5 Extraction of the human vote data

in search area, if high correlation value points exist, the average of the obtained motion vectors is used. Fig.2 shows an example of the detected motion data.

The motion data is expressed by θ and ρ as shown in fig.3, and the shifting quantity is voted to the θ - ρ space[6][7]. Fig.4 shows the voting result of the scene of fig.2. Here, the θ - ρ space is divided into the mesh beforehand, and θ is the angle of the direction of the motion vector. Using this angle, the value of the detected θ is always continuous and the progress direction of the human is preserved. The sign of ρ is used for distinction of two motion vectors, which are same on θ and ρ but differ in ρ angle (180 degrees each other). The acquisition of the voting data is repeated by increasing the input image number for a given period of time. Then the voted quantity and the continuance of the vote are examined at each position in the θ - ρ space. If the voted quantity and the continuance are satisfied enough, the voted quantity of the position is preserved. If not, the voted quantity of the position is eliminated. Consequently, the influence of the background change is decreased. By the way, the voting data obtained from passer-by is not concentrated on a point in the θ - ρ space. The vote data is, in fact, spread or shift a certain degree because of the clothes blow, body swing and so on. Therefore, the noticed vote position and its neighbor are examined. After the elimination of the vote data caused by disturbance, the existence of the human is detected in case that the residual vote data is confirmed in the θ - ρ space (fig.5). The residual vote data is called the human vote data.

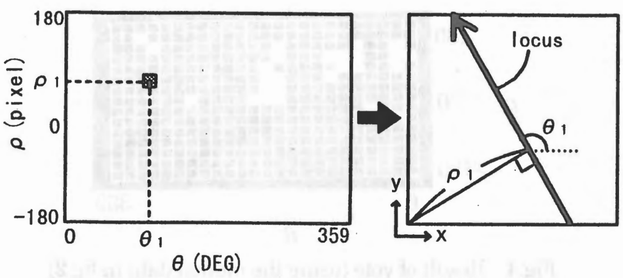


Fig.6 Path detection using inverse of θ-ρ transform

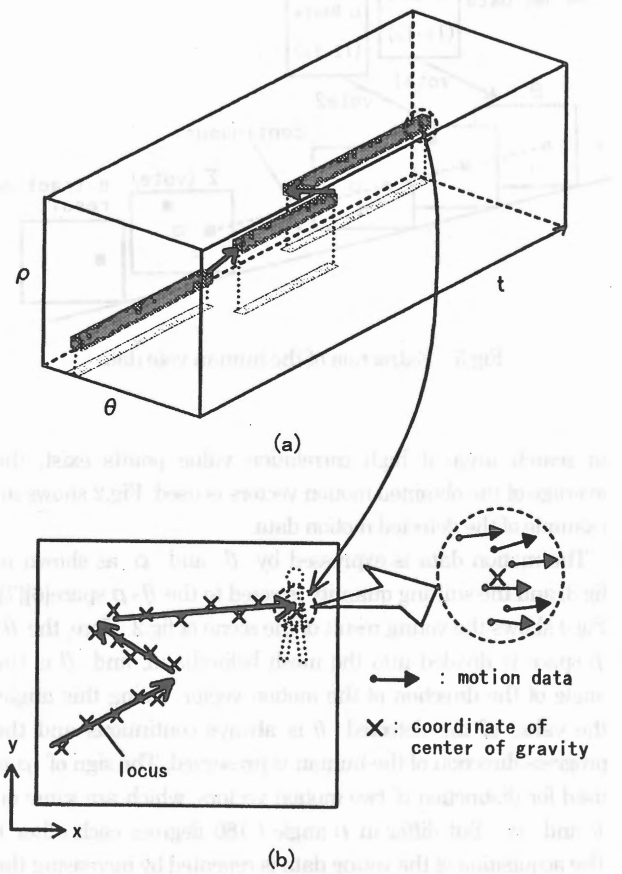


Fig.7 Locus detection using the extracted motion data. (a) θ-ρ-t space, (b)detection of line as a locus.

2.2 TRACKING

A straight line can be detected by computing the human vote data, using inverse of θ-ρ transform as shown in fig.6. This line detection method as a locus detection method has the advantage of simple algorithm and of reducing computation time. The detected line is significant enough for a locus of the straight line moving object. But, the starting position and the end position on movement are not detectable. The change of the course of the human is also undetected. Accordingly, the way using inverse of θ-ρ transform is adaptable to the scene of the passer-by with straight-line travel.

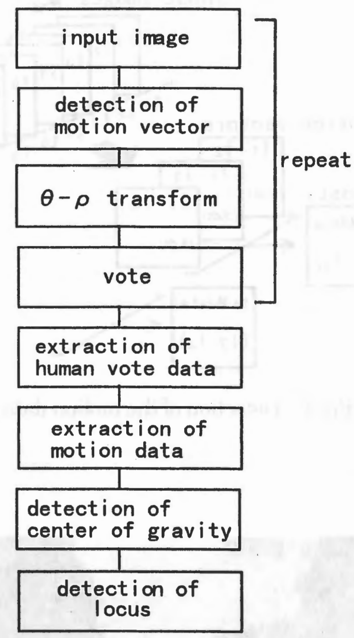


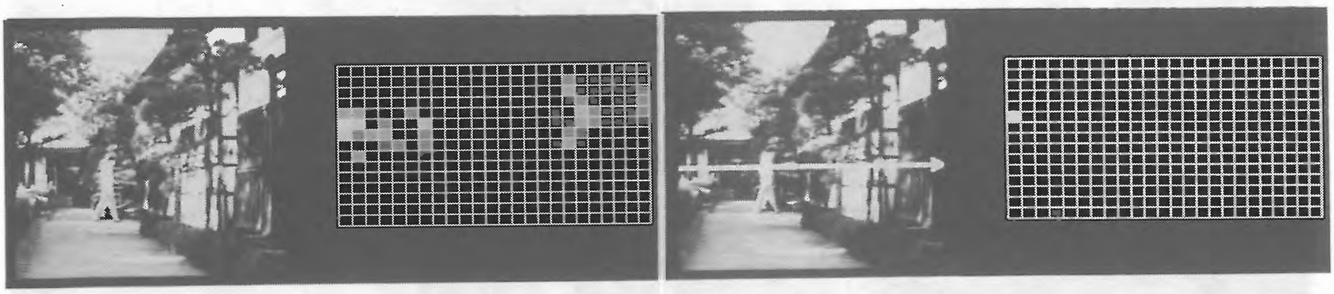
Fig.8 Flow of the processing

In order to judge the course change or the zigzag movement, the starting vote number and the end vote number on the human vote data are necessary. In the method, it is examined whether the vote is continuously performed at the position of the human vote data. If the vote is continuous, it is judged that the human keeps his course in succession. In case of the change of the course, the human vote data is represented in the θ-ρ-t space as shown in fig.7(a). And so, using the changing position (θ_i, ρ_i) and the number of t in the θ-ρ-t space, the change of the course can be detected. In tracking, to represent the human position, the motion data which contributes to the human vote data in the θ-ρ-t space is extracted. And the coordinates of center-gravity of the extracted motion data in one voting result is computed. Doing this computation about every voting results, the positions of the human in the image are detected. Then a line segment is computed by the least square approximation, using the obtained coordinates data. Consequently, the locus of the human is decided as shown in fig.7(b). Fig.8 shows the flow of the processing.

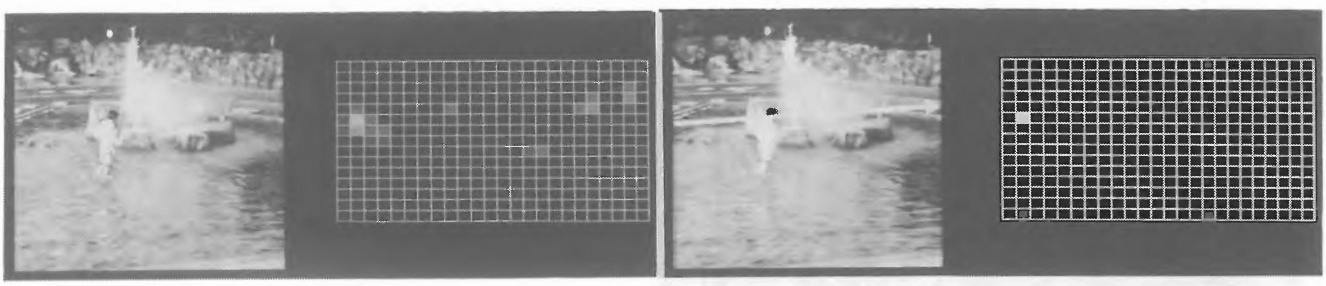
3. EXPERIMENTS AND DISCUSSION

The experiment was performed in a park. The time interval (Δt) of the input images is 0.5 second. The image is 128×128 (pixel) in size and is 256 in gradation. A template for computing the correlation is 3×3 (pixel) and the searching area is 21×21 (pixel). The resolution of the θ-ρ space is 24 (pixel) on ρ and is 15 (degree) on θ.

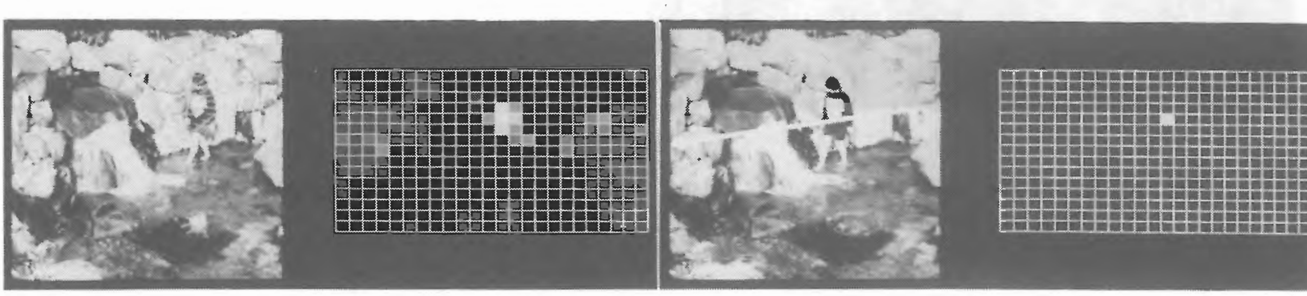
The examples of the experimental result are shown in fig.9 and fig.10. Fig.9 shows the locus detection result about the



(a) (b) (c) (d)

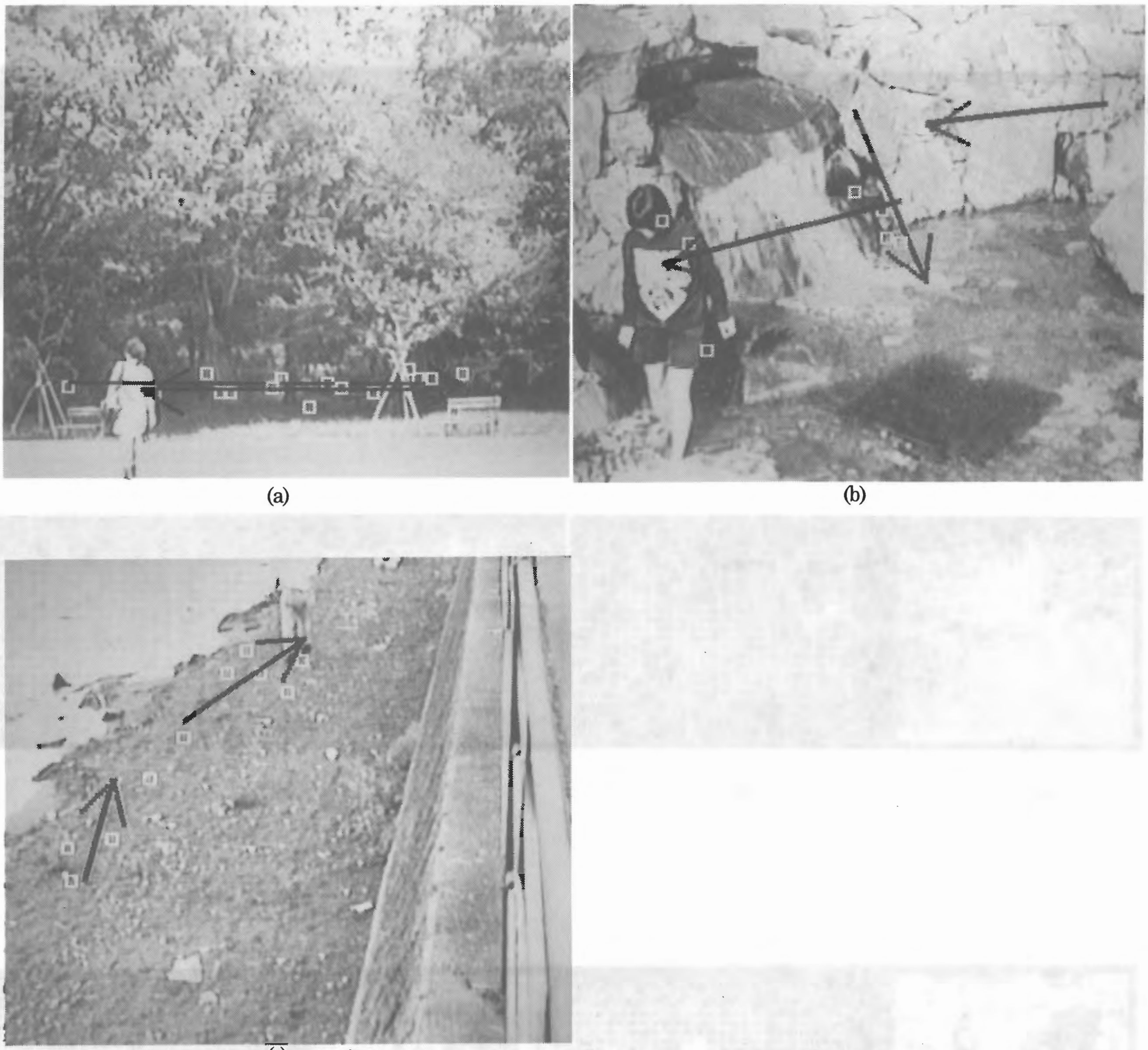


(e) (f) (g) (h)



(i) (j) (k) (l)

Fig.9 Experimental results (using inverse of θ - ρ transform). (a), (e), (i)input image and motion vector, (b),(f),(j)voting result, (c),(g),(k)locus detection result, (d),(h),(l)extraction result.



(a) (b) (c)

Fig.10 Experimental results (using the θ-ρ-t data). (a)thicket, (b)small waterfall and pond, (c)beach. Small square marks show the positions of the human.

scene of the human with straight-line travel. Here, the inverse of θ-ρ transform is used. Fig.10 shows the locus detection result in case of course change or zigzag movement. The background was as follows : a thicket in fig.9(a) and fig.10(a), a fountain and a pond in fig.9(e), a small waterfall and a pond in fig.9(i) and fig.10(b), a beach in fig.10(c). Small square marks in fig.10 show the positions of the existence of the human in the image.

The results of the human tracking were satisfactory approximately, though there was a little dislocation between the detection line and the true path of the human. In case that the human kept the straight line course, the locus detection results were good, using the inverse of θ-ρ transform (fig.9). In case of the scene of course change or zigzag progress, the human tracking results were also good, using the θ-ρ-t data (fig.10). The human vote data in the θ-ρ space was, in fact, spread or shift a certain degree because of the clothes blow, body swing and so on. Therefore, the dislocation of the line occurred. On the other hand, the human detection could not be influenced by the disturbance in many test scenes. However, there was the error in case of whole background severe sway (the branches, the leaves and their clear shadow on the road surface). The error caused because that the obtained vote data was similar to the human vote data. To eliminate this error, some improvement points are considered. They are the improvement of the peak detection accuracy in the θ-ρ space, the utilization of the transition of the extracted motion data and so on. About the detectable human speed, the method could be adapted to the scene that the extent of the speed is approximately from a walk with a stick to a quick pace.

As plans after this, the human tracking on the camera axis, the tracking of a curve movement and the plural passers-by tracking are given. About the human tracking on the camera axis, it is considered that the quantity concerned with the divergence and the convergence of the motion vector is useful. In case of the curve movement, voting result is not the concentrative on a local region, but is a pattern concerned with the curve movement in the θ-ρ space. Consequently, about the curve movement, it is considered that the improvement of the operation of the θ-ρ-t data is necessary. About plural passers-by, the addition of computing the θ-ρ-t data obtained by crossing and contact is necessary. By these improvement, the available extent of the method will be extended.

4. CONCLUSION

This paper described an outdoor human tracking method which is uninfluenced by a branches and leaves trembling in the wind, a light reflection on the surface of the water and so on, using a spatio-temporal information of the image. And the results of the experiment, which was performed to verify the effectiveness of the proposed method, were demonstrated.

In the method, the motion vector is obtained by a matching method, and the motion quantities are voted to a space which is represented by two parameters (the angle θ and the distance ρ). The existence of the human is detected by estimating the vote data which is concentrated on a local region in the θ-ρ space. Tracking the coordinates of center-gravity of the motion data which contributes to the human vote data, a locus of the movement of the human is detected. In the experiment at a park, the detection result was good by reducing the influence of the disturbance. The tracking result was approximately satisfied, though there was a little dislocation between the detected line and the true path of the human. Consequently, the efficacy of the method was verified. The future plan is as follows: (1) improvement on the operation of the θ-ρ-t data and the motion data for higher detection accuracy and higher tracking accuracy, (2)detection of the human various movement (direction, speed and curve), (3) plural passers-by tracking .

REFERENCES

[1] Hiroshi Yoshikawa, Hideto Fujiwara and Shin-ichi Kuroda : "Development of Video Surveillance System using Motion Information", Proc. of the 1991 IEICE Spring Conference, pp.7-178(D-466), 1991

[2] Atsushi NAGAI, Yoshinori KUNO and Yoshiaki SHIRAI : "Detection of Moving Objects in Changing Background", Trans. IEICE, Vol.J80-D-Ⅱ, No.5, pp.1086-1095, May, 1997

[3] Kazuhito MURAKAMI and Hiroyasu KOSHIMIZU : "A Method to Extract Convex Hull on θ-ρ Hough Transform Space", Trans. IEICE, Vol.J73-D-Ⅱ, No.4, pp.597-605, Apr., 1990

[4] Hideyoshi TOMINAGA, Naohisa KOMATSU, Takeshi MIYASHITA and Tsuyoshi HANAMURA : "A Motion Detection Method on Video Image by Using Hierarchical Pixels", Trans. IEICE, Vol.J72-D-Ⅱ, No.3, pp.395-403, Mar., 1989

[5] Satoshi SHIMADA and Fumio ADACHI : "Extraction of Moving Objects by Successive Matching", Trans. IEICE, Vol.J76-D-Ⅱ, No.10, pp.2196-2203, Oct., 1993

[6] N. Kobayashi and J. Yamaguchi : "Human Detection in Outdoors using Spatio-Temporal Information", Proc. of the 1st Symposium on Sensing via Image Information (SII'95), pp.39-42, 1995

[7] N. Kobayashi and J. Yamaguchi : "Human Tracking in Outdoors using Spatio-Temporal Information", Proc. of the 2nd Symposium on Sensing via Image Information (SII'96), pp.173-176, 1996