AUTOMATIC TRACK OF PEDESTRIAN FLOW IN THE CITY

Kayoko KITAI
Research Associate, Graduate school of Design
Kobe Design University
8-1-1 Gakuennishi-machi, Nishi-ku, Kobe,651-2196
E-mail: k-kitai@kobe-du.ac.jp
JAPAN

Commission V, Working Group IC V/III

KEYWORD: Pedestrians, Tracking, Pedestrian Flows, Occlusion

ABSTRACT

The purpose of this work is to analyze the movement of pedestrians in urban areas. At present there is not enough data to analyze the movement of pedestrians accurately.
In the architectural field, tracking and distribution data is now being done by hand. In the image processing field it is difficult to track the movement of pedestrians the same way as a car, because the original image will change and you cannot predict in which direction people will move.
It utilizes the algorithm system to show the trajectory of the image and which direction pedestrians will move, even though this movement can't be seen yet.
In this paper I report the problems to track automatically pedestrians using image processing. I develop a system to solve this problems, and the data surveys have proven that this method is effective.

1. PROBLEMS WITH TRACKING PEDESTRIANS

Using image processing, this is the point where two people or more meet each other on the street. Also it is the point where two or more meet, but the front pedestrian overlay others in frame when you record on videotape, then you can get only one when you extract the pedestrians. This is called " occlusion " part. We need to divide one in some pedestrians.
These is the method of matching frames to one anther for tracking pedestrians. If we use this method as the "occlusion " part, we will make mistake and we won't get each pedestrian's movement.
This paper propose the method of tracking pedestrians using the results of the segmentation of pedestrians with color and position informations in frame.

2. FLOW OF TRACKING PEDESTRIANS

Please look at Figure 1 to see how the Flow operates. The first part of this system is to extract the pedestrians from the original image, and the second part is to track the movement of these pedestrians.
I use the digital images converted by analog video images; RGB 256 colors and less than 600 pixels x 480 pixels.

2.1 EXTRACTING PEDESTRIANS

After making a survey of the way pedestrians move in town, with the sequence of moving images one can make a background image. I extract the pedestrians from the original image and calculate the difference between the original image and the background image, and label them so I can recognize each pedestrian. (Figure2)

2.2 TRACKING PEDESTRIANS

Extract pedestrians have five stages while they move the sequence of moving images.

a) Appearance :
Pedestrian appearances into sequence.
b) Movement :
Pedestrian moves in sequence.
c) Exit :
Pedestrian exits from sequence.
d) Attachment :
One block I extracted includes two or more pedestrians.
e) Separation :
One block I extracted separate each pedestrians.

I would like to suggest that you can track pedestrians after determining a correct stage from (a) - (e) stages using continuous two frames.
Pedestrian ordinary changes three stages, (a) → (b) → (c) in the sequence. In this case, you can match pedestrians which are in the frame next to each other, and by doing this you can track pedestrians which are exactly same.
In "occlusion" part, pedestrian has more two stages ,(d) attachment and (e) separation , beside (a) - (c) . I would like to suggest the following way to track pedestrians in this case.
Please look at Figure 3 to see how to track pedestrians including (d) attchment and (e) separation stages.

2.3 SEGMENTATION OF PEDESTRIANS

By using this Formula*, one can cluster and separate pedestrians with the K-mean method[1]), which is an easy way to gather data into one block.

$$dr_{k\_aver} = \sqrt{ks(x^2+y^2)+kc(r^2+g^2+b^2)} \qquad (1)$$

$: (ks, kc; 0 \le ks, kc \le 1)$

$: (x, y)$

$: (r, g, b; 0 \le r, g, b \le 255)$

This has been done for all of the five dimensions (R, G, B color spaces and x, y position spaces) .The extracted pedestrians have a hierarchical structure similar to the shape of a tree. (Figure 4) Each cluster has its own particular color and position data.

* K-mean method : It is one of the clustering methods. You can decide the representative points which you want to calculate, for example by making a grid. The you can reconstruct the clusters which you have already arranged according to your assumption.

2.4 MATCHING OF PEDESTRIANS

2.4.1 MATCHING OF FRAMES NEXT TO ONE ANOTHER

I would like to suggest that you can match clusters which are exactly the same, these results can be used for tracking pedestrians. (Figure 2)
I use the block matching method for matching pedestrians. Each cluster should be considered as one block. Those blocks which have the minimum differences in color distance are matched in the same cluster. In this way those clusters which are in the frame next to each other can be matched up. Those clusters which have the most similar color to each other are then matched together. This is called matching clusters.
Based on the results of matching up clusters of similar colors one can also match up pedestrians which are almost exactly the same. This is done by finding pedestrians which have the largest number of matching clusters. This is called matching PS_LABELS. (Figure 2 and Figure 5)
The pedestrians will not appear separately as in previous frames, so when tracking PS_LABELS you can only get one vector. At the end of the occlusion this is what happens because you are depending only on PS_LABELS with a large number of matching clusters.
Using matching PS_LABELS, you can distinguish between the "occlusion" part and other part.

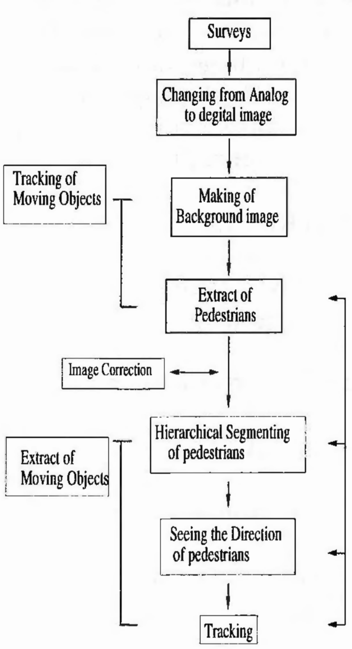


Figure1 Flow of Tracking Pedestrians

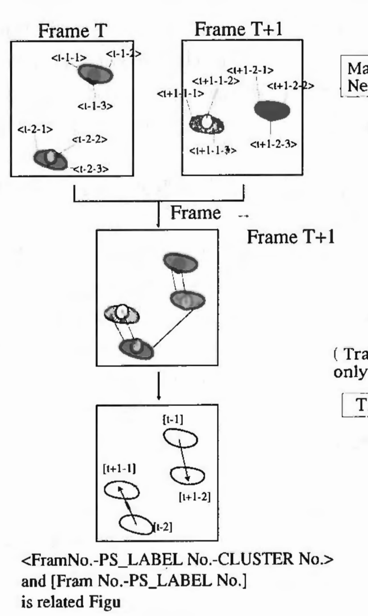


Figure2 Matching of Frames Next to One Another

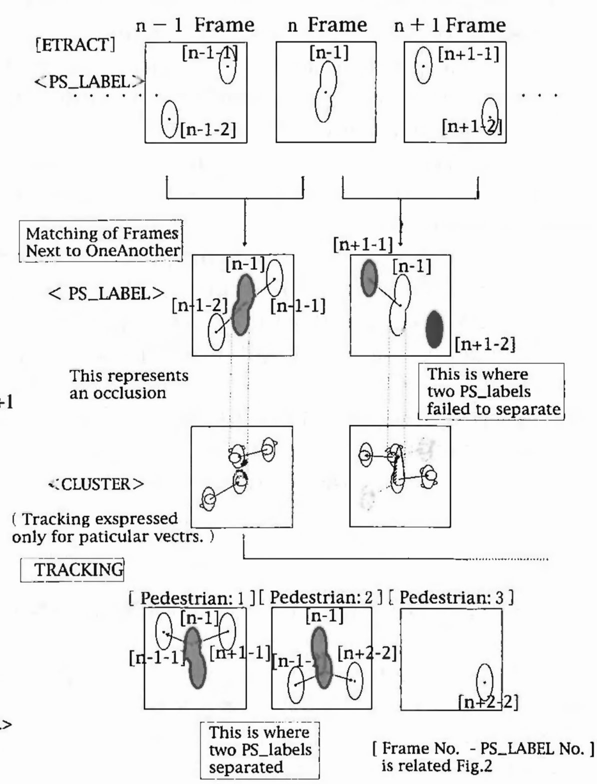


Figure 3 Tracking of pedestrians

2.4.2 MOVEMENTWITHTRACKINGPEDESTRIANS

One way to track pedestrians is by matching PS_LABELS from the first frame to the last frame. (Figure 3 and Figure 6) As you go along you can save matching data at each frame.

2.4.3 ATTACHMENT AND SEPARATION WITH TRACKING PEDESTRIANS (TRACKING OF OCCLUSION PARTS)

There is one problem with the "Occlusion" part . The PS_LABELS can be matched up with only one pedestrian at a time, so this part will divide one shape in some pedestrians. (Figure 3)
The clusters in the frame next to each other are connected to similar parts of the body. So you should try to find those clusters which are matched up and can separate attachment pedestrians.
Figure 7 shows an example of tracking max pixels. The cluster [0-P-1], which may have a maximum number of pixels, is one of the clusters in the pedestrian [0-P]. The following is the result of tracking cluster [0-P-1] in frames 0 to 2: [0-P-1]→ [0-Q-4]→ [0-Z-1]. The result of tracking pedestrian [0-P] is shown on the right side of Figure 7.

3 EXPERIMENTS USING FLOW

Look at Figure 8

This shows an example , especially [(d) attachment and (e) separation] part ,of the tracking method which I suggest.
This is the result of tracking pedestrians and matching PS_LABELS between 32 and 35 frames. In these images, <33-2> and <33-3> are connected <34-2> and <34-2> separated <35-1> and <35-3>. But matching PS_LABELS failed : <34-2> could not separate <35-1> and <35-3>. <34-2> matched up <35-3> and <35-1> was shown as a new pedestrian ((a) appearance).
Tracking with matching clusters in the "occlusion" part was successful : <34-2> could separate <35-1> and <35-3> and the trajectory of pedestrians <32-2> → <33-2> → <34-2> → <35-1> and <32-3> <33-3> → <34-2> → <25-3> were got.
[psn_no.16] was not the occlusion.
[psn_no.17] and [psn_no.18] are connected between 24 frame and 25, and separated between 31 frame and 32.
Figure 8 show an example of how this was done in Ginza.

4 CONCLUSION

This contribution should be considered important because it utilizes the flow to track pedestrians including that they cross each other and the distance among some pedestrians is very close. How this is done will be explained in the following.

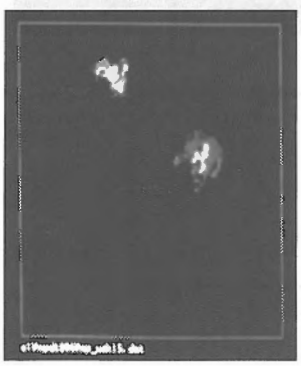
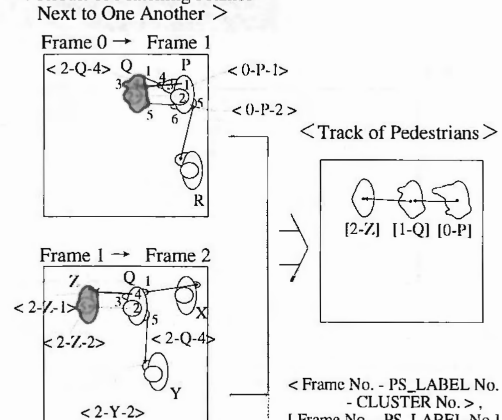


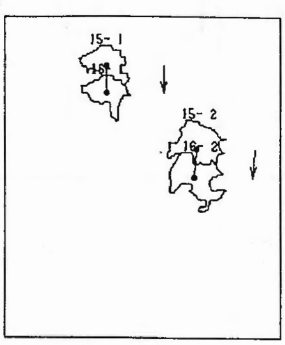
Figure4 Extract of Pedestrians

Figure5 Matching of Frames Next to One Another

< Result of Matching Frames Next to One Another >



< Frame No. - PS_LABEL No. - CLUSTER No. > ,
[Frame No. - PS_LABEL No.] is related Fig.7.



e:¥nwp_labl¥np_usk15. lbl
e:¥nwp_labl¥np_usk16. lbl

Figure6 Tracking of PS_LABES

Frame 0 ⇒ *Frame 1*

Pedestrian P (5 8 0 *pixels*) :
* [0 − P 1] → [1 − Q 4]
(3 5 0 *pixels*) (2 8 0 *pixels*)
[0 − P 2] (1 2 5) → [1 − Q 2] (1 5 5)
[0 − P 3] (5 5) *has no matching clusters.*
[0 − P 4] (2 0) → [1 − Q 1] (1 0)
[0 − P 5] (1 5) → [1 − R 2] (5 0)
[0 − P 6] (1 5) → [1 − Q 5] (3 0)

Pedestrian Q (5 0 0 *pixels*) :
[1 − Q 1] (1 0) → [2 − X 4] (1 0)
[1 − Q 2] (1 5 5) → [2 − Z 2] (8 0)
[1 − Q 3] (2 5) → [2 − Z 5] (5 0)
* [1 − Q 4] (2 8 0) → [2 − Z 1] (3 2 0)
[1 − Q 5] (3 0) → [2 − Y 2] (5 0)

Figure7 Tracking Pedestrians (Occlusion Part)

The Cluster *[0-P1]* has the maximum number of pixels in *Pedestrian P*. The result of matching Cluster*[0-P1]* in *Frame 0* corresponds with the Cluster *[1-Q4]* in the next frame. *Pedestrian P* in *Frame 0* corresponds with *Pedestrian Q* in *Frame1*. Regarding *Pedestrian Q* from *Frame 1* to *2* you match clusters the same way as you did in the previous frames. The trajectory of *Pedestrian P* in *Frame 0* corresponds with the following, which is the result of matching clusters from *Frame 0* to *2*. *Pedestrian P => Pedestrian Q => Pedestrian Z*

First , after making a survey of the way pedestrians move in town with the sequence of moving images ,one can make a background image. I extract the mobile object from the original image and calculate the difference between the original image and the background image, and label them so I can recognize each object.
Second , I take the labeled pedestrian (PS_LABEL) and segment it with the cluster method in five dimensions, R, G, B color spaces and x,y position spaces. The moving objects have a hierarchical structure similar to the shape of a tree.
Finally , we match the frames with the clusters and the labeled objects. Based on the results of this matching one can track the moving objects. Using this method it is also possible to track the point at which some pedestrians meet each other on the street.

REFERENCE

Kitai, K.,1997. Automatic Tracking System of Pedestrians Using Image Processing Method., Jour. of Architecture, Planning and Environmental Engineering, pp.195-200. (in Japanese)

Kitai,K.,1996. A Study of Automatically Tracking Pedestrian Flow ., Proc. of MVA'96., pp.257-260., 1996 .

Tsujimoto,M. Shida,K. and Tatebe,K. 1992. Application of Digital Image Processing Techniques to Analysis of Pedestrians Behavior., JAPEE no.436, AIJ, June. pp.41 - 47 . (in Japanese)

Izumi,N. Morikawa,H. and Harashima,H. 1991. Combining Color and Spatial Information for Segmentation. 1991 SNCR D-680, IEICE.,pp.7-392. (in Japanese)

Sato,A. Mase,K. and Suenaga,N. 1992.An Object Extraction Using Background Model on RGB Color Space.1992 SNCR D-438, IEICE., pp. 7-180. (in Japanese)

Original Image

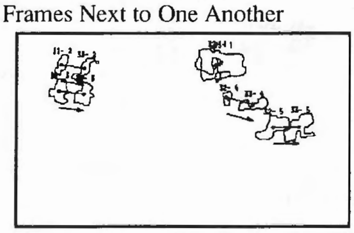
GNZ_32

GNZ_34

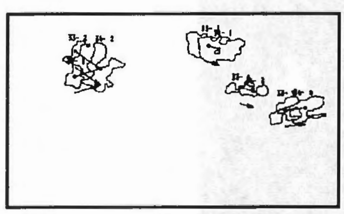
GNZ_33

GNZ_35

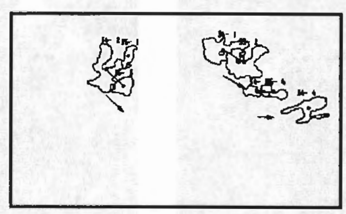
Matching of Frames Next to One Another



GNZ_32→ GNZ_33



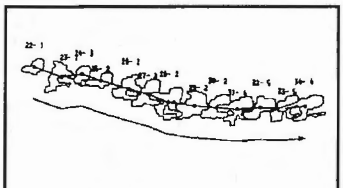
GNZ_33→GNZ_34



GNZ_34→GNZ_35

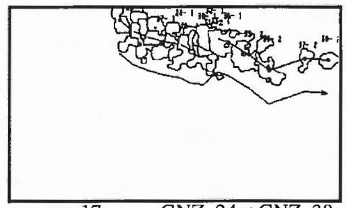
Tracking of Pedestrians

(b) Movement

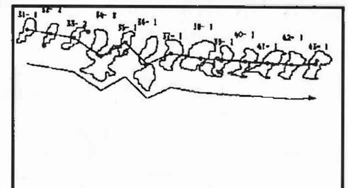


psn_no. 16 GNZ_22→GNZ_34

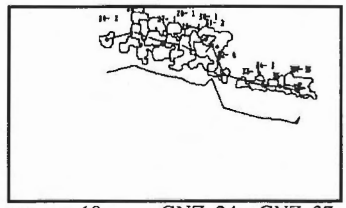
(d)Attachment and (e)Separation



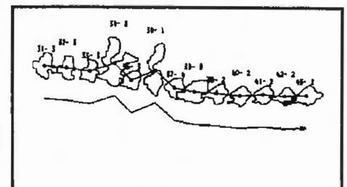
psn_no. 17 GNZ_24→GNZ_38



psn_no. 19 GNZ_31→GNZ_43



psn_no. 18 GNZ_24→GNZ_37



psn_no. 20 GNZ_31→GNZ_43

Date : 1995.10.28 (sat.) 2:00 ~ 3:00p.m.
Place : Ginza, Chuo-ku, Tokyo
VideoCamera : SONY CCD-TR705

This experiment in Ginza primarily shows two people walking together and the distance between two people is very close. There are 25 people of the sequence in Ginza.

[Frame No. — PS_LABEL No.] is related Fig.8.

Figure 8 Survey on The Street