MVA '96 IAPR Workshop on Machine Vision Applications, November. 12-14, 1996, Tokyo, Japan

A Study of Automatically Tracking Pedestrian Flow

Kayoko Kitai*
Graduate School
Kobe Design University

Abstract

The purpose of this work is to analyze the movement of pedestrians in urban areas.[1] At present there is not enough data to analyze the movement of pedestrians accurately.[2]

In the architectural field, tracking and distribution data is now being done by hand.[3] 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.[4]-[6]

It utilizes the algorithm system to show the trajectory of the image and which direction those objects (pedestrians) will move,even though this movement can't be seen yet.

The data surveys have proven that this method is effective.

1 Automatic Tracking Algorithm

Please look at Fig.1 to see how the Algorithm operates.

The first part of this system is to extract the moving objects from the original image, and the second part is to track the movement of these objects.

2-1. Extracting Moving Objects

After making a survey of the way pedestrians move in town, with the sequence of moving images one can make a background image(1). I extract the mobile object(2) 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. (Fig.2)

2-2. Tracking Moving Objects

1) Segmentation of moving objects

By using this Formula (1), one can cluster and separate moving objects with the K-mean method, which is an easy way to gather data into one block. (Fig.3)

$$dr_{k\_aver} = \sqrt{ks(x^2+y^2)+kc(r^2+g^2+b^2)} \qquad (1)$$
$$(0 \le ks, kc \le 1)$$
$$(0 \le r, g, b \le 255)$$

(x, y) represents exact location of each pixel.
(r, g, b) represents each value.

This has been done for all of the 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.

* Adress : 8-1-1 Gakuennishi-machi, Nishi-ku, Kobe, Hyogo ,651-21, Japan. E-mail : kayoko@da2.so-net.or.jp

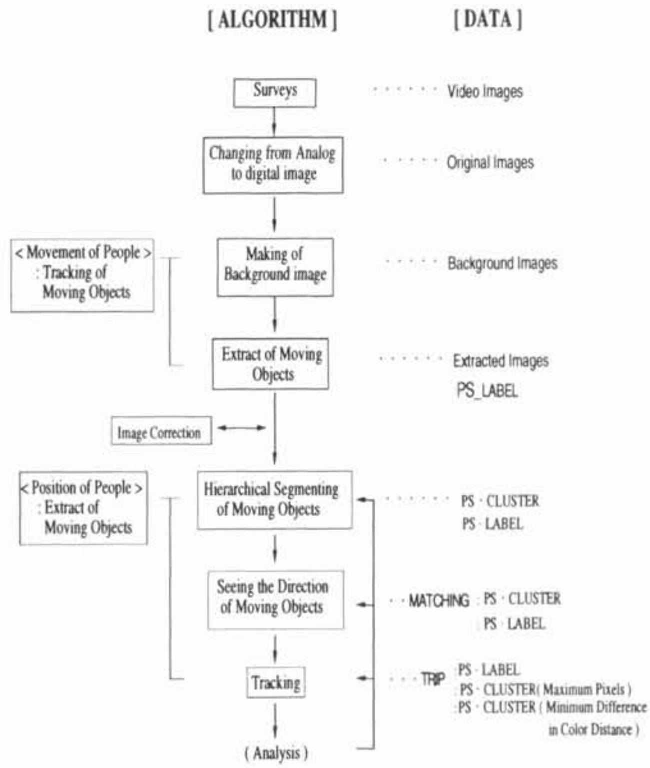


Fig.1 Algorithm of Automatic Tracking

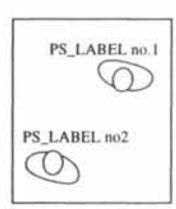
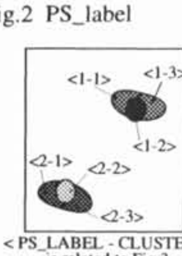


Fig.2 PS_label



< PS_LABEL - CLUSTER >
is related to Fig.3.
Fig.3 Segmentation of Moving Objects

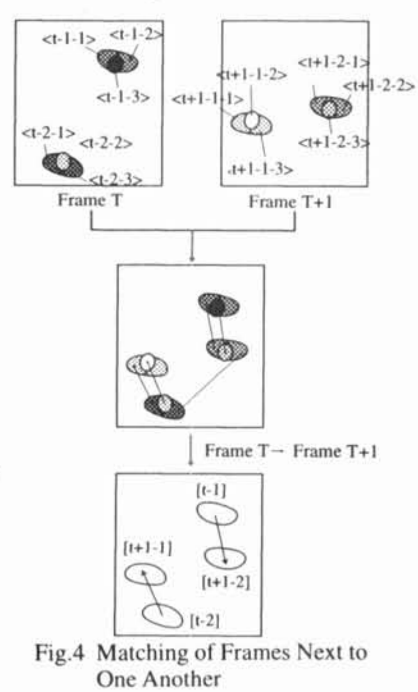


Fig.4 Matching of Frames Next to One Another

Each cluster has its own particular color and position data.

2) Matching of Moving Objects

I would like to suggest that you can match clusters which are in the frame next to each other, and by doing this you can match moving objects which are exactly the same. (Fig.4) Also I would like to suggest that these results can be used for tracking moving objects.

I use the block matching method , which is one way to match moving objects. 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 moving objects which are almost exactly the same. This is done by finding objects which have the largest number of matching clusters. This is called matching PS labels.

3) Tracking of Moving Objects

I would like to suggest three ways to track moving objects which have the largest number of matching clusters. There are three types of data which one is targeting. You can select the most suitable way to track original images.

3-1) Tracking PS Labels

One way to track moving objects is by matching PS labels from the first frame to the last frame. (Fig.5) As you go along you can save matching data at each frame. There is one problem with this method. The PS labels can be matched up with only one object at a time, so the "Occlusion" part will divide one shape in some objects. (Fig.6) For that reason you might make a mistake by using this method for tracking moving objects.

3-2) Tracking of Occlusion Parts

There are two ways of doing this. (Fig.7, Fig.8) 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. This can be done by choosing those clusters which have the largest number of pixels. The other way is by choosing clusters which have minimum difference in color distances. When there is minimum time ddifferency, data for clusters will not change very much.

The tracking method for Fig.4 thru 7 are combined in Fig.8.

2 Experiments Using Algorithms

Look at Fig.9.

This shows an example of the tracking methods which I suggest.

I would like to tell you some things I have learned from these experiments.

You should try tracking with PS Labels. You can substitute another way for the frames which do not work. The record tracking data will select the most suitable way for you.

(i) For anything other than the Occlusion parts, tracking with PS Labels is the best way because of two factors, pro-

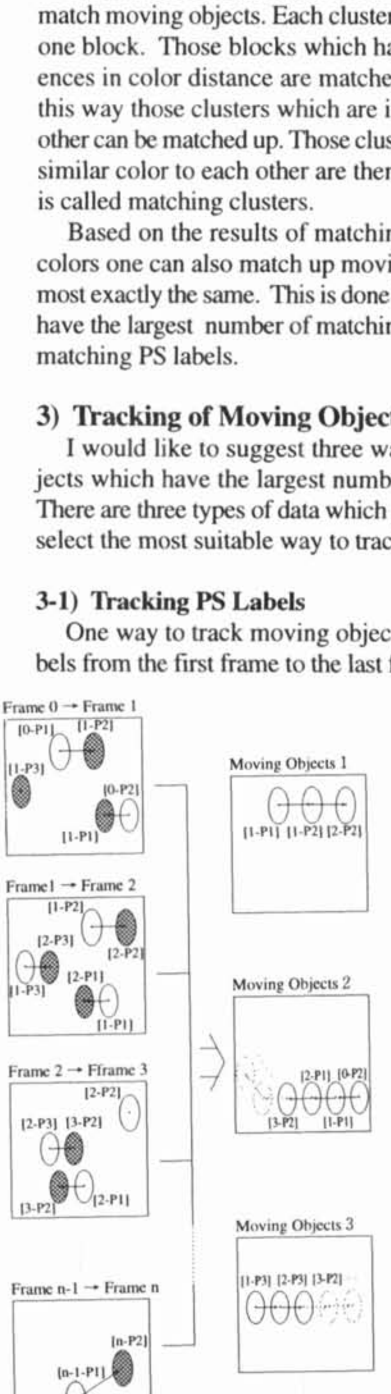


Fig.5 Tracking PS_labels

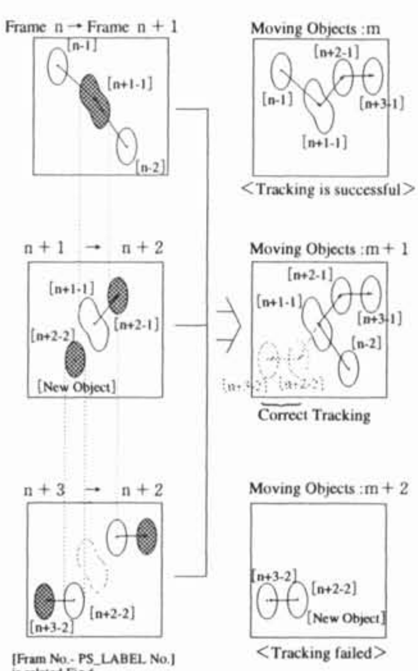


Fig.6 Problems with Tracking PS_labels

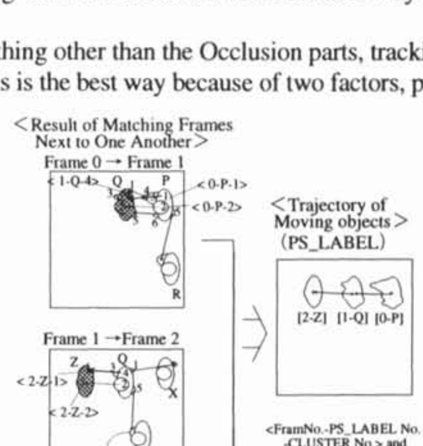


Fig.7 Tracking of Clusters

In Fig.6 the same moving objects are shown separately as [m+1] and [m+2] because [m+2] is shown as a new object. The moving objects 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. When the moving objects separate only one vector will appear and other moving objects can't be tracked.

Fig.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 moving object [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 moving object [0-p] is shown on the right side of Fig. 7.

If you are tracking the occlusion part the same way as in Fig. 7, you can track moving objects successfully like in Fig. 8, including the occlusion part.

cessing speed and accuracy.

(ii) Regarding Occlusion parts the best way is tracking with matching clusters. When there is a big difference in color distance you should use the clusters with the most pixels. When the image has many gaps and is not complete you should use clusters which have minimum difference in color distances.

Fig.10 shows an example of how this was done in Ginza.

3 Conclusion

This contribution should be considered important because it utilizes the algorithm system to show the trajectory of the image and which direction those objects (pedestrians) will move,even though this movement can't be seen completely.

I would like to explain what resulted from designing the algorithm system.

1) The trajectory of objects (pedestrians). Using this method it is also possible to track the point at which some pedestrians meet each other on the street. This is an example of an occlusion.

2) The location of where objects (pedestrians) are in an area. This is equal to the extracted image.

3) It uses a color image, but the usual method of image processing is binary.

4) The benefit of this system is that only the moving object can be seen in color. The processing speed can be made faster because there is a smaller amount of information.

This system can be used to help athletes, particularly runners and ball players to improve their movements, by analyzing the extracted moving image.

Appendices

1) Background Image : I took the average or mode for each pixel in all frames.

2) Extract of Moving Object : I extracted the moving objects using the background image, because I could get a clearer boundary of the extracted objects than by using sequential differential extraction.

References

[1] K. Kitai : " Automatic Tracking System of Pedestrians Using Image Processing Method", Ph D Thesis, Tokyo University, March, 1996 (in Japanese)

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

[3] Y. Naka : " A Study on Complicated Passenger Flow Railway Station", Ph D Thesis, Tokyo University, March, 1977. (in Japanese)

[4] S. Takeuchi and K. Nagata :" Method of Extraction of Moving Object from Images" ,Tchnical Resarch Report of Shimizu Co. vol 56, Oct. 1992, pp.79~86(in Japanese)

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

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

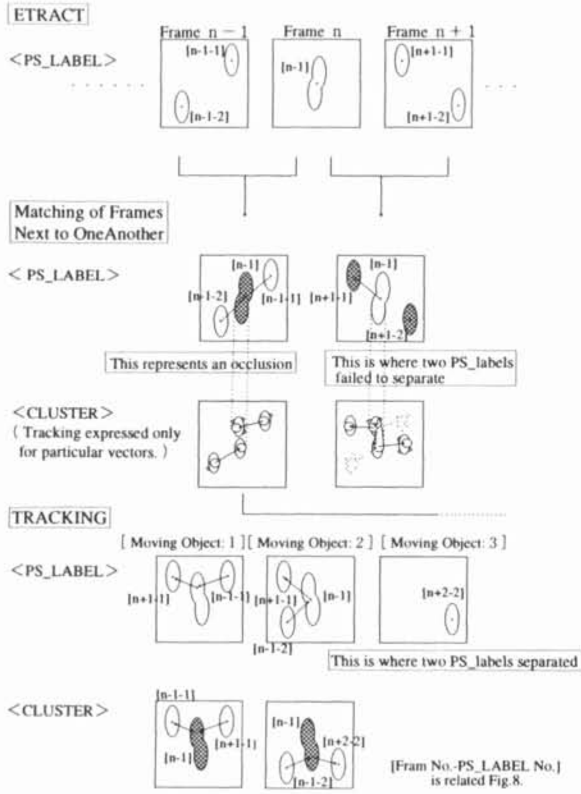
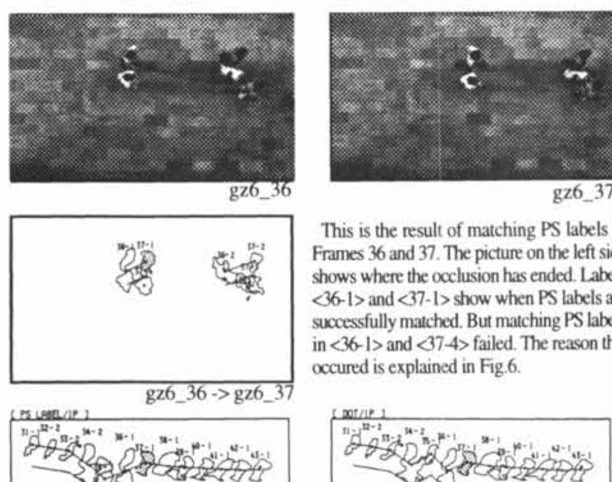
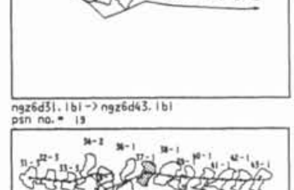
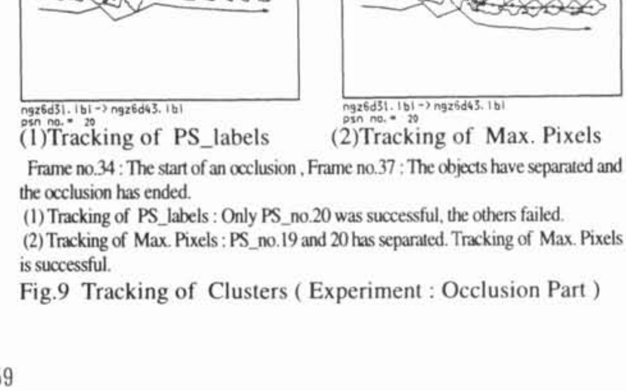


Fig.8 Tracking of Moving Objects (Included Occlusion Part)





This is the result of matching PS labels in Frames 36 and 37. The picture on the left side shows where the occlusion has ended. Labels <36-1> and <37-1> show when PS labels are successfully matched. But matching PS labels in <36-1> and <37-4> failed. The reason this occured is explained in Fig.6.



(1)Tracking of PS_labels (2)Tracking of Max. Pixels

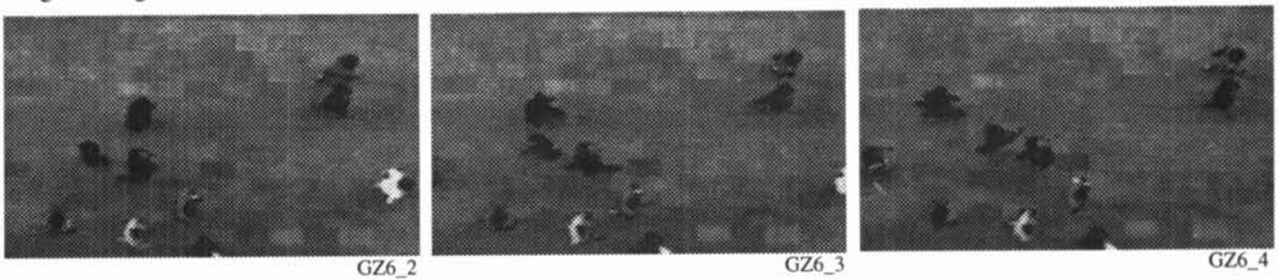
Frame no.34 : The start of an occlusion , Frame no.37 : The objects have separated and the occlusion has ended.

(1) Tracking of PS_labels : Only PS_no.20 was successful, the others failed.

(2) Tracking of Max. Pixels : PS_no.19 and 20 has separated. Tracking of Max. Pixels is successful.

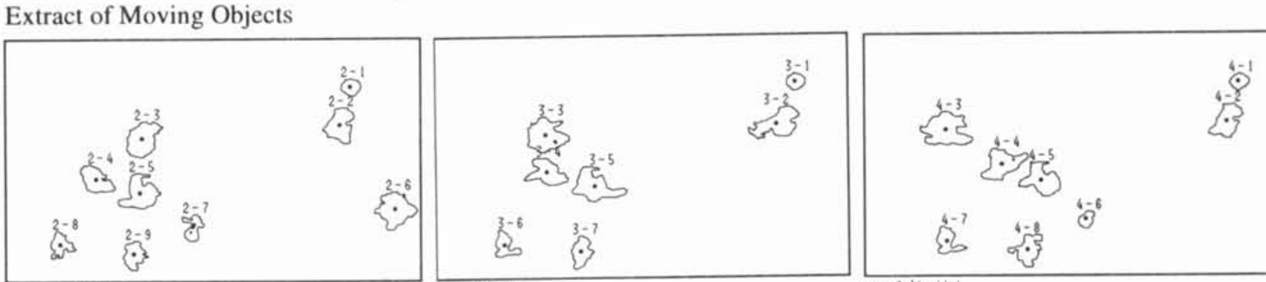
Fig.9 Tracking of Clusters (Experiment : Occlusion Part)

Orignal Images



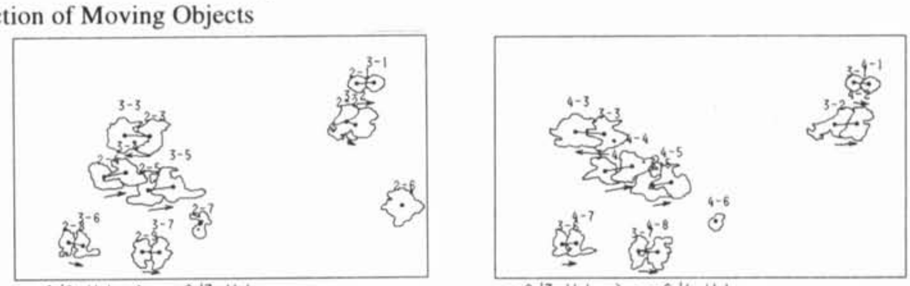
GZ6_2 GZ6_3 GZ6_4

Extract of Moving Objects



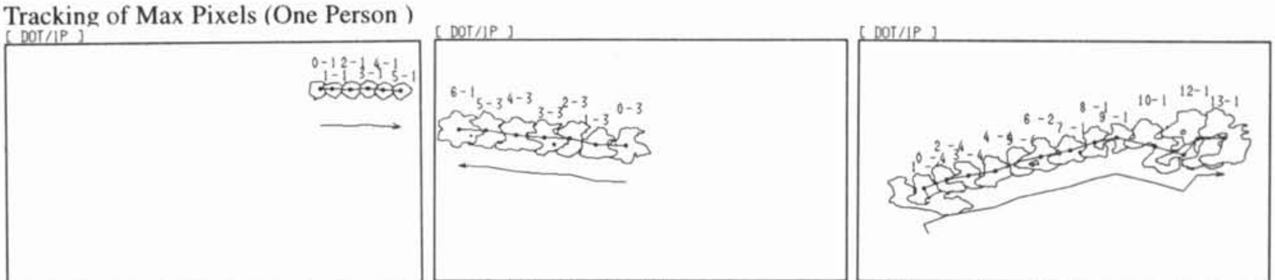
ngz6d2. lbl ngz6d3. lbl ngz6d4. lbl

Seeing the Direction of Moving Objects



ngz6d2. lbl -> ngz6d3. lbl ngz6d3. lbl -> ngz6d4. lbl

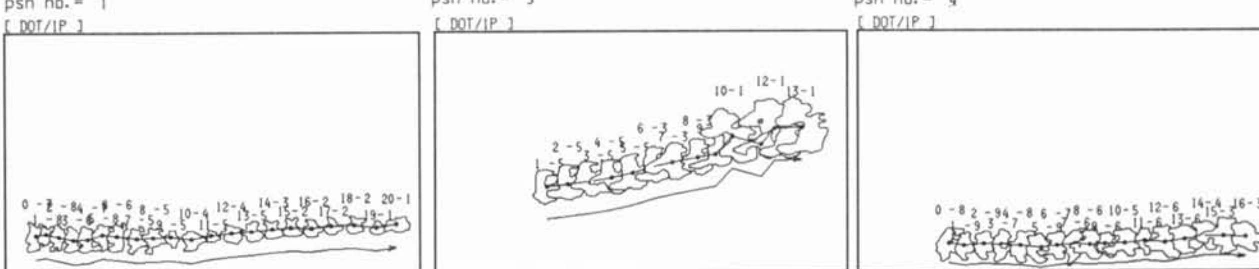
Tracking of Max Pixels (One Person)



ngz6d0. lbl -> ngz6d5. lbl
psn no. = 1

ngz6d0. lbl -> ngz6d6. lbl
psn no. = 3

ngz6d0. lbl -> ngz6d13. lbl
psn no. = 4

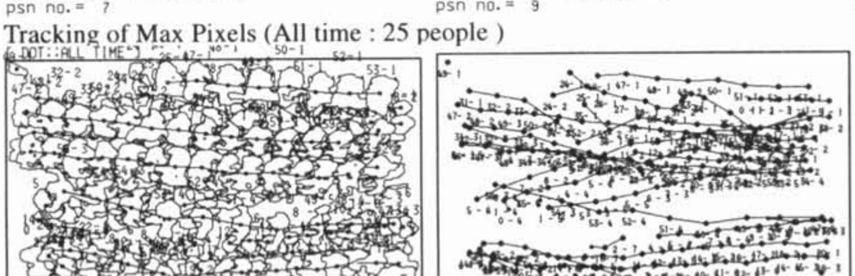


ngz6d0. lbl -> ngz6d20. lbl
psn no. = 7

ngz6d1. lbl -> ngz6d13. lbl
psn no. = 9

ngz6d0. lbl -> ngz6d16. lbl
psn no. = 8

Tracking of Max Pixels (All time : 25 people)



ALL_TIME(objects)

ngz6d0. lbl -> ngz6d59. lbl
psn no. = 25

This experiment in Ginza primarily shows two people walking together and the distance between two people is very close. These extracts show that there are many examples of occlusions taking place. Rare cases of tracking individuals are unsuccessful because you cannot separate moving objects in the occlusion part.

In Ginza I had the most success in tracking pedestrian flow by tracking maximum pixels. Here is a good example of how to track pedestrian flow.

Fig.10 Experiment on The Street (Ginza)