Objects Tracking in Image Sequence by Optimization of a Penalty Function

Akio Sakata, Hiroyuki Shimai, Kazuyuki Hiraoka, and Taketoshi Mishima
Department of Information and Computer Sciences, Saitama University
255 Shimo-Okubo, Saitama, Saitama, 338-8570, JAPAN
Phone: +81-48-858-3723, FAX: +81-48-858-3723
E-mail: sakata@me.ics.saitama-u.ac.jp

Abstract: We suggest a novel approach to the tracking of multiple moving objects in image sequence. The tracking of multiple moving objects include some complex problems(crossing (occluding), entering, disappearing, joining, and dividing) for objects identifying. Our method can settle these problems by optimization of a penalty function and movement prediction. It is executable in real time processing (more than 30 Hz) because it is computed by only location data.

1. Purpose

Tracking multiple moving objects from image sequences captured by a fixed camera is very important topic for many video applications (visual surveillance system, visual interface for human-computer interaction, etc). There are a number of reports about this topics in recent years [1][2]. If moving objects are tracked, the system can easily pay attention to them and can recognize or check it.

Objects tracking task consists of two steps, detecting objects and identifying each detected objects. Detecting objects is easy because the camera is fixed in this case. As for the most simple and typical approach, we can detect objects by the background subtraction[3]. However, the identifying each object is not so easy because special cases (crossing (occluding), entering, disappearing, joining, and dividing) may exist in the scene. Moreover, complex combinations of these cases can occur when there are more than three objects in one scene: this is a very difficult problem.

To solve this problem we suggest objects identifying method by optimization of a penalty function and by movement prediction. This penalty function consists of the impossibility of fundamental five cases and is computed only the location of objects (center of gravity). The movement prediction is done by linear prediction using historical data of the center of gravity when the occlusion is occurred.

This method enable us to deal with the complex identifying problem as a simple optimization problem and a simple prediction problem. Then, it is executable in real time processing (more than 30 Hz) because it is computed by only location data.

We describe the identification method in Section 2. Experiments are discussed in Section 3 and conclusions are in Section4.

2. Methods

To achieve objects detection, we use the background subtraction and the painting algorithm. They are one of the popular and basic approach of object detection. They can extract the number of objects and the center of gravity of each objects. Then, we can identify each objects based on these informations. The flowchart of proposed identification algorithm is shown in Figure 1.

2.1 Creating initial proposal of the correspondence to each objects

The first stage of the algorithm is the creating initial proposal of the correspondence to each objects between successive image frames. The object in current frame is assoceated with one object in previous frame using the distance measures between each objects in successive frames. This proposal is correct in many cases because we suppose that difference of time is small between two frames. However, there is the case that this proposal is wrong. For example, as shown in Figure 2, it is the wrong correspondence to two objects using the distance measures (a solid arrow). In this case, we needs making amendment proposal for it (a dotted arrow).

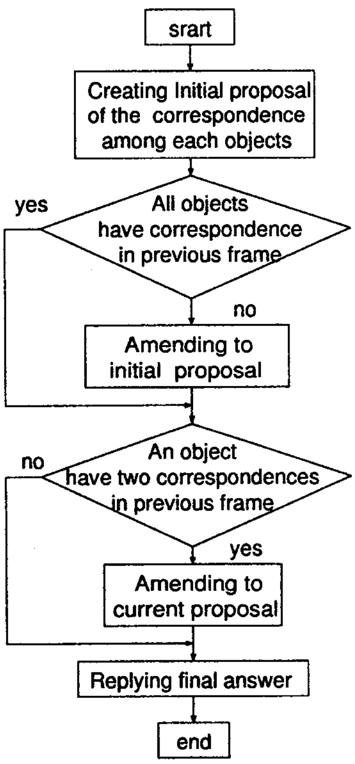


Figure 1. Identification Algorithm

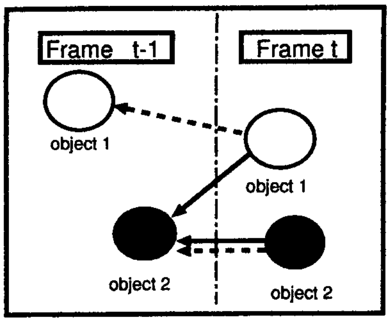


Figure 2. Definition of Temporary Label

2.2 Amending to initial proposal

Amending process is based on the criterion by penalty function. The penalty function is defined to find the reliable proposal of correspondence to each objects between two successive image frames. It express the impossibility in fundamental five cases (transfer, joining, dividing, intruding, vanishing).

We intend to amend to the proposal for optimizing the penalty function. However, thorough optimization is very difficult because complex correspondance cases can occur. Furthermore, it is disadvantage for real time processing because computation cost is very large. Then we approximate a reliable proposal by taking into account only following two cases. First case is that there is the object in previous frame which is associated with no object. Second case is that there is the object previous frame which is associated with more than two objects.

Figure 3 shows the example of first case. In this case, object 1 is associated with no object in initial proposal. Then we can think about three cases of pattern, which shows in Figure 4.

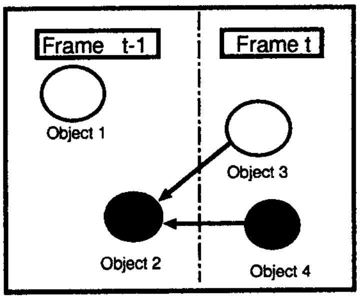
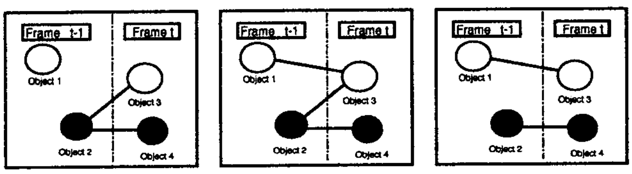


Figure 3. Case of Example 1



pattern 1 pattern 2 pattern 3

Figure 4. Case of Pattern

Pattern 1 reveals that object 1 disappears in current frame and object 2 divide into object 3 and object 4 in current frame. Pattern 2 reveals that object 1 and one divided object 2 unite and it is set to object 3 in current frame and another divided object 2 set to object 4 in current frame. Pattern 3 reveals that object 1 moves to where object 3 is in current frame and object 2 moves to where object 4 is in current frame. In these pattern, we calculate penalty function about each case and choose the case which have minimum value among them.

Figure 5 shows the example of second case. In this case, object 1 is associated with two objects in current frame. Then we can think about four cases of pattern, which shows in Figure 6.

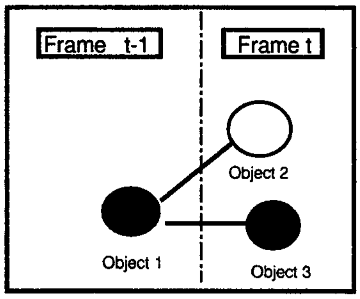


Figure 5. Case of Example 2

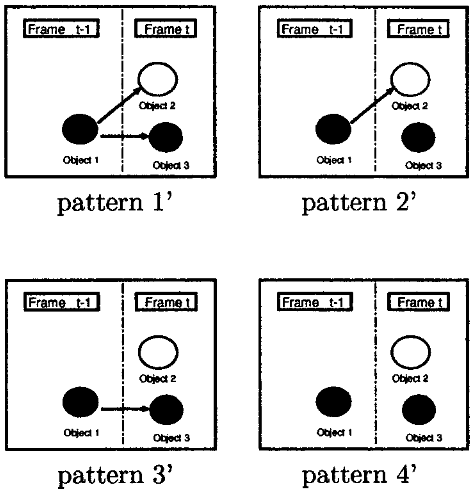


Figure 6. Cases of Pattern 2

Pattern 1' reveals that object 1 is divided object 2 and object 3 in current frame. Pattern 2' reveals that object 1 moves to where object 2 is in current frame and object 3 appears in current frame. Pattern 3' reveals that object 1 moves to where object 3 is in current frame and object 2 appears in current frame. Pattern 4' reveals that object 1 disappears and object 2 and object 3 appear in current frame. We calculate penalty function about each pattern and choose the case which have minimum value among them.

2.3 Penalty Function

The penalty function of a proposal is expressed as follows.

Penalty function

$$R = Rt + Rj + Rd + Ri + Rv$$

- The transfer penalty function
  $R_t = (1/1100) \times$ (a square moving distance)

- The joining penalty function
  $Rj = 3.0 \times$ (sum total of the moving distance of all objects in joining)

- The dividing penalty function
  (One object)
  $R_{f1} = 9.0 \times$ (sum total of the moving distance of all objects in dividing)
  (One object overlapping by two objects)
  $R_{f2} = 4.0 \times$ (sum total of the moving distance of all objects in dividing)

- The intruding penalty function
  (Near the border of the image(within 2 pixels))
  $R_{i1} = 5.0 \times$ (the number of appearing objects)
  (Far from the border of the image)
  $R_{i2} = 10.0 \times$ (the number of appearing objects)

- The vanishing penalty function
  (Near the border of the image(within 2 pixels))
  $R_{v1} = 1.5 \times$ (the number of vanishing objects)
  (Far from the border of the image)
  $R_{v2} = 5.0 \times$ (the number of vanishing objects)

Note that parameters showing here are in the case that the image size is 180x120 pixels.

2.4 Movement prediction

When moving object is crossed (occluded) by other objects, these objects are merged and separated after several frames. In this case, current labels of each objects have to correspond to former ones. Then, the movement of two objects can be estimated by a linear prediction using the historical data of the center of gravity.

2.5 The feature of process

All these processes, we use only two information, center of gravity and object label. Then, we have two merit for real time application. One is that the computation time is decreased significantly than using another information such as shape, color, texture, etc. Another is that the history of information can be recorded because these information is small.

3. Experiments

We have performed two experiments on multiple objects tracking. All objects were extracted by background subtraction. A desktop PC (Pentium4 1.8GHz ,DDR SDRAM PC2100 256MB) was used for these experiments and the program have been developed on LINUX.

3.1 The artificial image sequences of five objects

We have performed experiments on multiple objects tracking for artificial image sequences. In this sequence, five moving objects are included and they caused fundamental five cases(crossing (occluding), entering, disappearing, joining, and dividing). These images are 24 bit color images and the size of image is 720x480. The several frames of this sequences are shown in Figure 7. Figure 8 shows the tracking result of five objects. It shows that each object are tracked exactly notwithstanding complex situation is occurred.

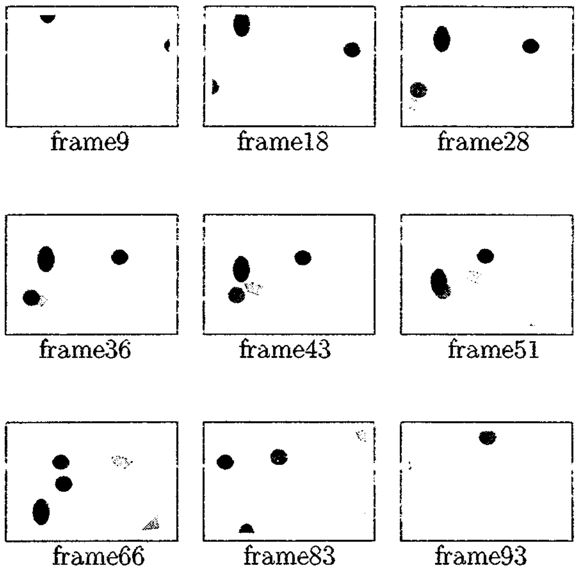


Figure 7. image sequence for tracking experiments 1

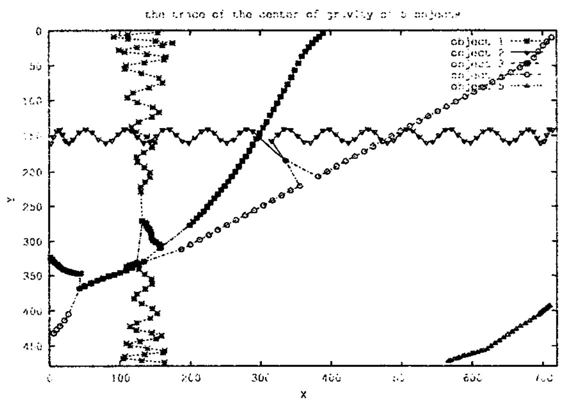


Figure 8. Tracking result of multiple moving objects

3.2 The real image sequences of two people

We have performed experiments on two persons tracking for real image sequences. These images are 24 bit color images and the size of image is 180x120. The several frames of this sequences are shown in Figure 9. The results of tracking each person on every image sequences was presented by adding the tracking position. Black and white dots show the position of each person when they move independent, and gray dots show the position of persons when they are crossing(occluding) and form one segment (moving object). Figure 10 shows the tracking result of two persons. It shows that our methods can track precisely for real image sequences.

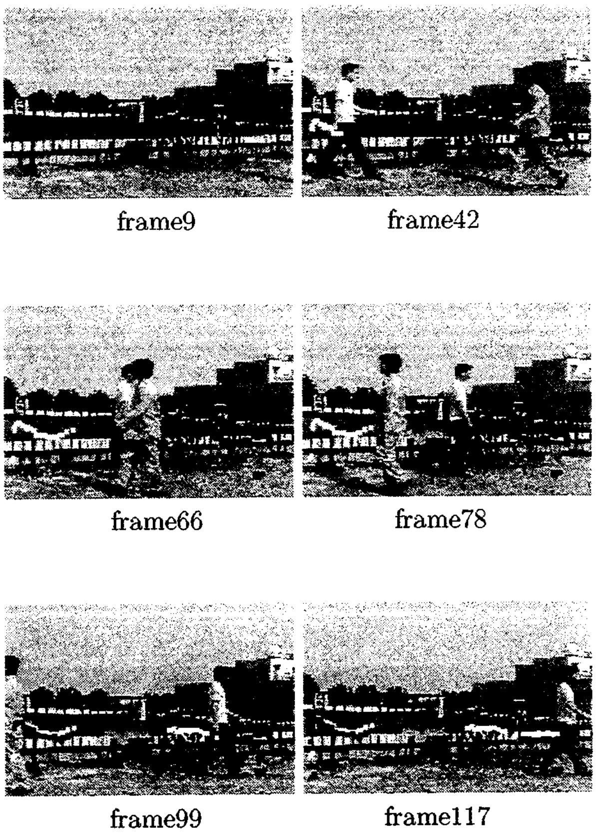


Figure 9. image sequence for tracking experiments 2

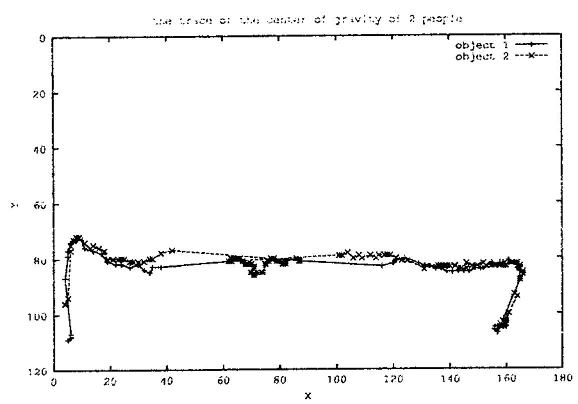


Figure 10. Tracking result of two moving people

Table 1 shows the processing speed of this method. Though all process (detection by background subtraction and painting algorithm, identifying by proposed algorithm) is executed on the single PC, all the processing can be done in 293.2 frame per second when the image size is 180x120. It is about ten times faster than real time processing (30Hz).

4. Conclusion

We suggested objects identifying method by optimization of a penalty function and by prediction of objects movement for multiple objects tracking. The proposed technique could reduce the difficulty of dealing with the correspondence to each objects between successive image frames. Moreover the processing speed of this method could execute about 10 times faster than real time processing, if the image size is 180x120. It means that other technique can be added for more accurate tracking and higher process (estimation or understanding of object's action, recognition, etc).

Table 1. average of the processing speed

image size	average of the processing speed(frame/sec)
720x480	18.41
180x120	293.2

Acknowledgement

This work has been partly supported by JSPS (Japan Society for the Promotion of Science), 14580405.

References

[1] S. A. Niyogi and E. H. Adelson, "Analyzing and Recognizing Walking figures in XYT, " Proc. *IEEE Conf. Computer Vision and Pattern Recognition*, pp. 469-474, 1994.

[2] N. Johnson, and D. Hogg, "Learning the distribution of object trajectories for event recognition, " Proc. *Image and Vision Computing(IVC)*, 14(8), pp. 609-615, 1996

[3] H. Shimai, T. Mishima, T. Kurita,and S. Umeyama, "Adaptive background estimation from image sequence by on-line M-estimation and its application to detection of moving objects, " Proc. *Infotech Oulu Workshop on Real-Time Image Sequence Analysis(RISA)*, pp. 99-108, 2000.