

Research on Video Analysis and Tracking Technology in Swimming Video

Ni An

*Physical Education Department, Northeast Agricultural University,
Harbin 150001, China
491687484@qq.com*

Abstract

This paper proposes the tracking method for swimmer in the video with Kalman filtering prediction. The Kalman initiates the swimmer with feature such as the location of object and the size of its enclosing rectangle which can be got in the detection proceeding, which are used to predict the feature of the tracked swimmer in subsequent frames and ascertain the scope for searching the tracked object. The results of experiments show that this method can be used to provide information for controlling cameras and this can help to track the swimmer in the next frames.

Keywords: *Swimming, Detection of moving object, Tracking of Moving object, Kalman filter*

1. Introduction

The aim of moving object tracking system based on videos is:

(1) To decide when a new object will enter the observation scope of the system and to initialize its motion model [1-2];

(2) To calculate the association between the moving object detected through existing frames and current tracked object. Under the premise that object motion is stable, it's possible to track it in videos. However, in real scene, it's very difficult to extract features of tracked targets, because:

(1) Features like noises, light variation, reflection and chaos can make object's characteristics obscure;

(2) Multi-moving targets exist simultaneously, especially when multiple objects with similar features shelter from each other, their characteristics can be mixed;

(3) Features of objects with non-rigid links can often vary;

(4) Specific requirements in real use are different [3].

After detection of moving region and extraction of targets in video sequence, object extraction becomes source data of object tracking [4-5]. In the tracking process, the first step is to generate object's feature template; then search matching in the prediction range of the preceding frame of the object, update feature template of objects in the best matched moving area and correct the position of objects [6-7]. In the meantime of predicting the position of object in next frame. To be brief, tracking means finding out object's exact moving region in one-frame image [8-9].

The moving object tracking based on videos is a very important and hot research topic in the field of computer vision, digital video and image processing and pattern recognition [10]. The moving object tracking in video sequence means, on the basis of detecting moving objects in video sequence, utilizing object's effective features and proper matching algorithms to find the position of candidate targets similar to target templates [11]. Since moving object tracking is a significant link between moving object detection and the behavior analysis and comprehension. In real use, it not only provides object's motion trail and accurate location, and also helps detecting moving objects.

2. Technology for Tracking Moving Objects in Video Sequences

The tracking of motion objects in video image sequences is to make consecutive or time-sampling discrete measurement of object's location in image. Through estimation of object moving behavior, the location of objects at next moment in the image can be predicted and the same moving object at different frames in image sequence can be connected to get the complete movement of each motion object. Time sampling of tracking technology usually takes two trains of thought: (1) tracking through object identification; (2). tracking through motion detection. The first method determines the location of target through the investigated object recognized from each frame image. This method includes object recognition and matching. The second method discovers and decides the location of moving object through detection of it. Without considering object's shape and size, the method can detect any object.

For tracking method of moving objects in video sequence, whether rigid or non-rigid object, single or multiple objects, whether the object is captured by still or moving camera, whether the object in complicated or simple scenario, it needs to choose one or more features which can only represent the object according to it and its environment, then search from following images the location of object which matches best with the chosen feature. The study of object tracking algorithm involves two aspects: (1) how to withdraw effectively object feature; (2) how to make rapid and accurate search and matching in the current image as per known object feature. So far, there're lots of object tracking algorithms. Their common characteristic is increasing the search speed and accuracy degree of object matching with the selected tracking feature. Hence, the core issue of object tracking includes two points: a. how to search quickly the location of object feature; b. how to make accurate match of feature template.

3. Method for Tracking Swimming Athletes

With regards to the key technology for tracking athletes in swimming video sequence analysis, the paper presents object feature extraction method based on Blob analysis and target prediction tracking method based on Kalman filter [11].

Blob analysis refers to analyzing the connected domain of identical pixels in images. Such domain is called Blob. Blob analysis needs to solve mainly problems of Blob statistics and information fetch. Blob statistics is to sum up the number of Blob which suffices the condition in images; tag each Blob so that it's possible to make attributive statistics of each relative Blob to such tag. The information fetch is to get all information necessary for Blob analysis, *i.e.* aggregating, which is the most important step for object feature fetch, and also most time-consuming part. It requires to optimize the algorithm several times.

During the scanning of connected region, it needs to obtain a few geometric features of it, including its boundary point of line segment, minimum bounding rectangle, area, perimeter and centroid position *etc.*

In measuring the size of target region, its area $S(\bullet)$ can be used as a measuring parameter. For the region $R(x, y)$, $S(\bullet)$ is defined as pixel number in the region.

$$S(R_i(x, y)) = \sum_{(x,y) \in R_i(x,y)} f(x, y) \quad (1)$$

The other parameter to measure target region feature is perimeter, referred to as $P(\bullet)$; it's defined as the number of pixel dots on regional boundaries.

$$P(R_i(x, y)) = \text{Number of pixels } (x, y) \text{ on the edge of the boundary} \quad (2)$$

You can also calculate the centroid of target unit, its expression is $R_i(x_0, y_0)$, where (x_0, y_0) is:

$$\begin{cases} x_0 = M_{10}(R_i(x, y) / M_{00}(R_i(x, y))) \\ y_0 = M_{01}(R_i(x, y) / M_{00}(R_i(x, y))) \end{cases} \quad (3)$$

Apart from the above information, the information about Blob boundaries is required, including profile of boundary points and location of each extreme point up and down, left and right. It's only Blob is calibrated well that accurate integration can be done and target position marked.

At some time, it's no need to concern the specific representation of chain code of boundary; instead, to concern some key points of it. With those points, people can determine a so-called bounding rectangle of Blob. The rectangle can provide the following information:

$$\begin{aligned} \mathbf{top} &= \min_y \{(x, y) \mid (x, y) \in R_i(x, y)\} \\ \mathbf{bottom} &= \max_y \{(x, y) \mid (x, y) \in R_i(x, y)\} \\ \mathbf{left} &= \min_x \{(x, y) \mid (x, y) \in R_i(x, y)\} \\ \mathbf{right} &= \max_x \{(x, y) \mid (x, y) \in R_i(x, y)\} \end{aligned} \quad (4)$$

All kinds of information can be stored in the Blob of the information structure of the body.

After Blob information got from images, we need to aggregate Blobs which belong to the same moving object based on the feature of each Blob, which is also called region merging, to realize extraction of target information, thereby to complete the transition from Blob information to target information, offering necessary feature information to track further targets.

Region combination utilizes Blob's distance and area to verify whether multiple Blob targets belong to the one region and then merge those that meet conditions. During the merging, it firstly makes initial screening of Blob in searching area, excluding Blobs with the area below threshold out of consideration, to remove noises. This step is named Blob area filter. In this case, region merging is a kind of image processing method for eliminating the interference of noises. Region margining depends on conditions like centroid location, upper and lower boundary condition of bounding rectangle. With those conditions, it's possible to merge Blob meeting conditions and having similarities as one target; meanwhile, such information as size of target, feature of bounding rectangle frame and target's histogram can be acquired through Blob's information.

While merging Blob, we can compare each Blob's information with target. If Blob's one information can help change target information during the merging of Blob to the target, it needs to replace original target information with new one. Simultaneously, the centroid of target is closely related with that of each Blob. During the Blob merging, the centroid is updated gradually. When the entire merging process completes, information of the whole detected object is obtained, such as centroid, area, feature of bounding rectangle block, target's pixel grey histogram. Besides, information of object feature can be added or reduced as per the needs of following tracking modules; then, use such information to form feature vectors of target and store in the structural body of target information.

During shooting and tracking, in order to control precisely the camera, it's necessary to locate the target in camera's field of view. So we can consider fetching the minimal bounding rectangle block in the area to which moving object belongs as output.

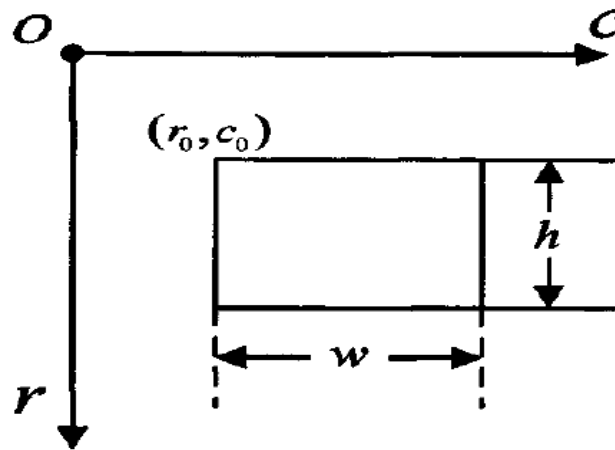


Figure 1. The Rectangular Frame in the Image Coordinate System

Image coordinate system and bounding rectangle block are shown in Figure 1, where the original point locates in the upper left corner of the image. In the experiment, the bounding rectangle block vector is defined like:

$$z = (r_0, c_0, h, w)^T \quad (5)$$

Where, r_0 represents the upper left corner of the rectangle frame in the image coordinates. c_0 represents the column coordinates of the upper left corner of the outer rectangle frame in the image. The width and height of the rectangular frame are represented by h and w .

4. Simulation Experiment on Swimmer Target Detection and Feature Extraction Based on the Information Fusion of Time-Space Domain

We conducted simulation test regarding swimming video sequence. The original video sequence is collected from www.swim.ee, with image resolution 352*288 pixel, being used to fetch grey images at the 56th to 57th frame of the sequence. It is show in Figure2.

First of all, we introduce moving target detection approach based on information fusion of time-space domain and use it to process the sequence. It is show in Figure3.

Then, get a binary image with self-adaptive thresholding segmentation; further, to confirm automatically if motion objects got through that method have feature information of players, by analyzing connected component identification and Blob of those binary images. It is shown in Table1

Table 1. The Information of the Players Outside the Rectangular Frame in the Swimming Video

Frame number	r_0 (Testing)	c_0 (Testing)	h (Testing)	w (Testing)
56 th Frame	102	35	155	210
57 th Frame	102	45	156	208
58 th Frame	146	32	110	230
59 th Frame	35	22	98	241
60 th Frame	155	24	89	168
61 th Frame	144	45	76	156
62 th Frame	145	23	76	178
63 th Frame	147	23	65	209
64 th Frame	154	37	67	201
65 th Frame	154	36	71	189

66 th Frame	156	40	78	178
67 th Frame	150	42	76	167

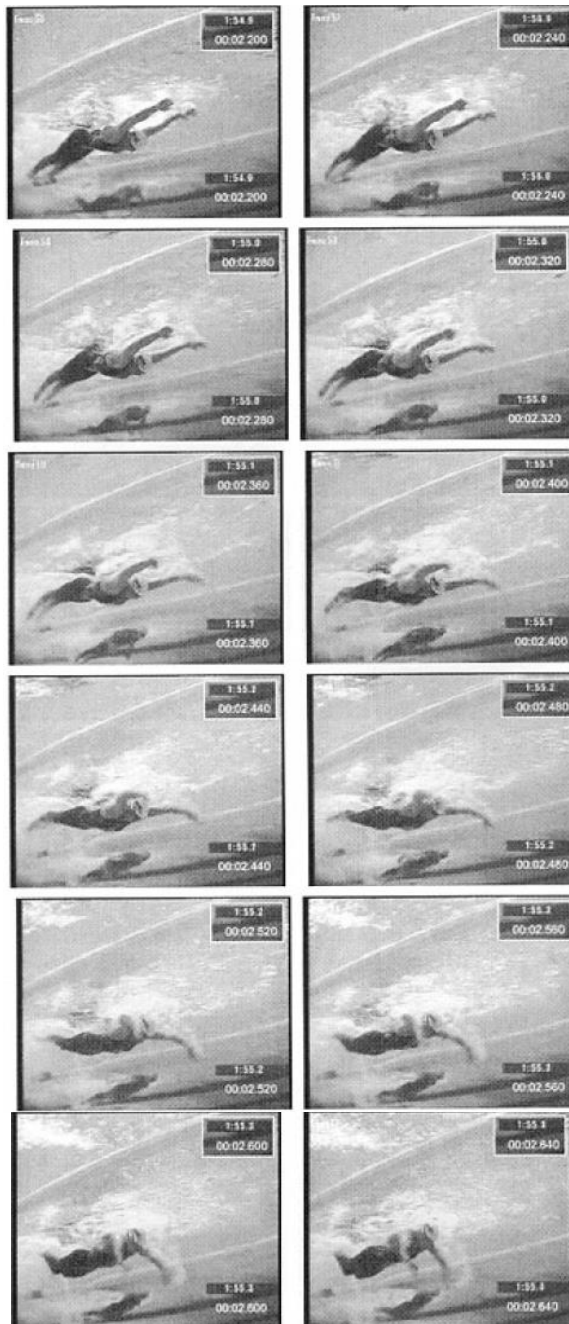


Figure 1. Gray Scale Images of the Original Video Sequences (from Left to Right, from Top to Bottom, 56 to 67 Frames)

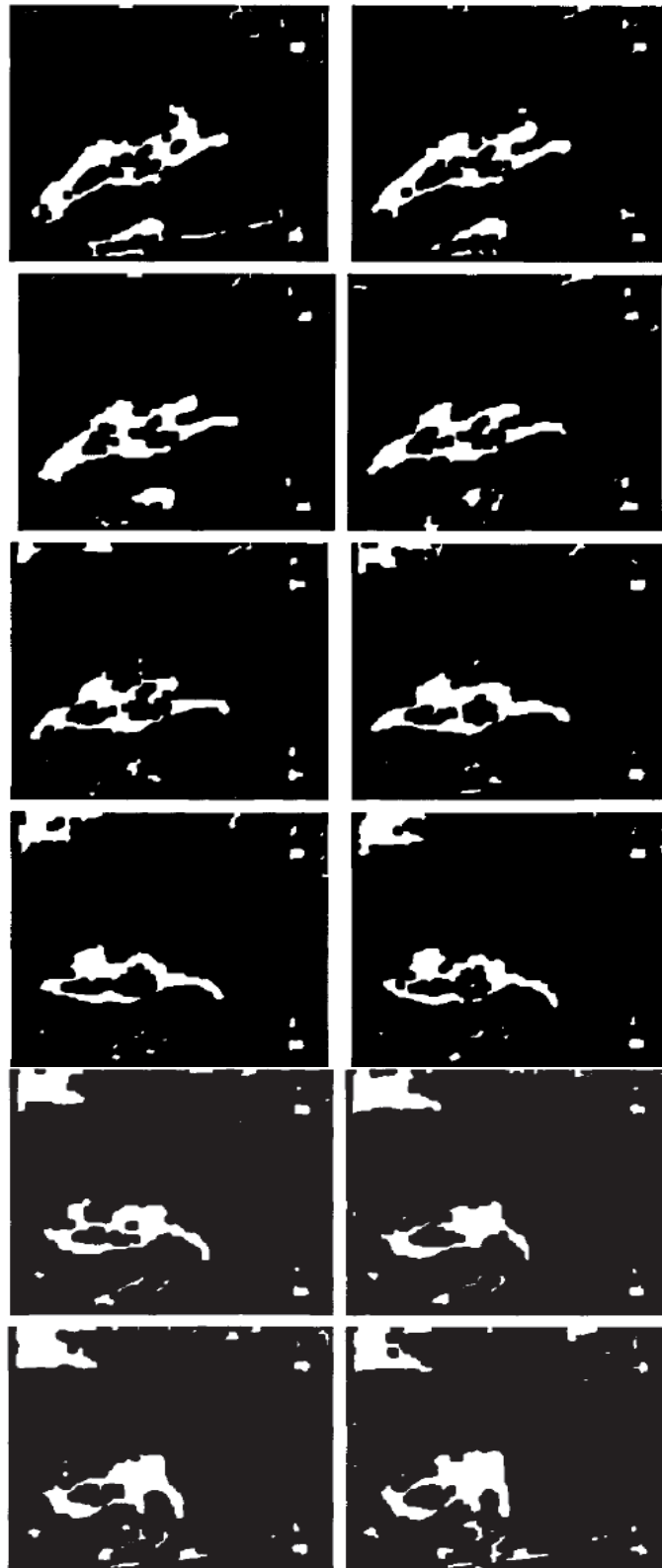


Figure 2. The Two Value Image of the Moving Object Detection and the Adaptive Segmentation of Figure 1(from Left to Right, from Top to Bottom, 56 to 67 Frames)

5. Predictive Tracking Principle Based on Kalman Filter

Kalman filtering sticks to the basic idea: to update iteratively by forecasting and renewing two steps in time sequence to reduce errors caused by system noise and observation noise as to get the best status parameter.

Since the interested moving target changes along with time, so the description of target feature should be dynamic accordingly. In temporal sense, moving object falls behind the matching foreground object by one frame. So before matching, it needs to predict its feature at next frame according to its historical moving records. During target tracking, we use Kalman filtering method to anticipate the location of targets and use it as the starting point of the whole matching process, in order to decrease dramatically iteration times and enhance computing speed [12-13].

Steps of Kalman filter used for tracking moving objects

Initialization: it's the first time that initialized filter is utilized; x initial value is given object detection position and object size; likewise, we can take into account the initialized state covariance;

Prediction: with Kalman filter, we can foresee the relative position of object; regard location x_k as searching center to search the object;

Correction: locate the object and find x_k through Kalman filter; then correct its state based on its actual location.

6. Kalman Filtering Prediction Simulated Experiment on Swimming Videos

After system model and observational model are defined, Kalman filter is employed to forecast the minimum bounding rectangle block information including swimming players.

Assume swimmers make linear normalized movements. The system state can be expressed with six-dimension vectors including location coordinate of bounding rectangle block of upper left corner in images, height and width of the rectangle block and change rate of height and width. Therefore, the system state vector and transition matrix can be defined as:

$$x = \begin{bmatrix} r \\ c \\ \Delta h \\ \Delta w \\ h \\ w \end{bmatrix}, \quad F_{k+1} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

With the prediction tracking method based on Kalman filtering, simulation test was performed for swimming players. The video sequence demarcated by bounding rectangle block at the 56th to 57th frame was extracted. It is shown in Figure3.

By means of information fusion based on time-space domain and Blob analysis, the paper fulfilled the purpose of swimming object detection and its feature extraction. The minimum bounding rectangle block information which contained swimmers was acquired. Then with Kalman filter to predict tracking, the bounding rectangle information which included swimmers was obtained. It is shown in Table 2.

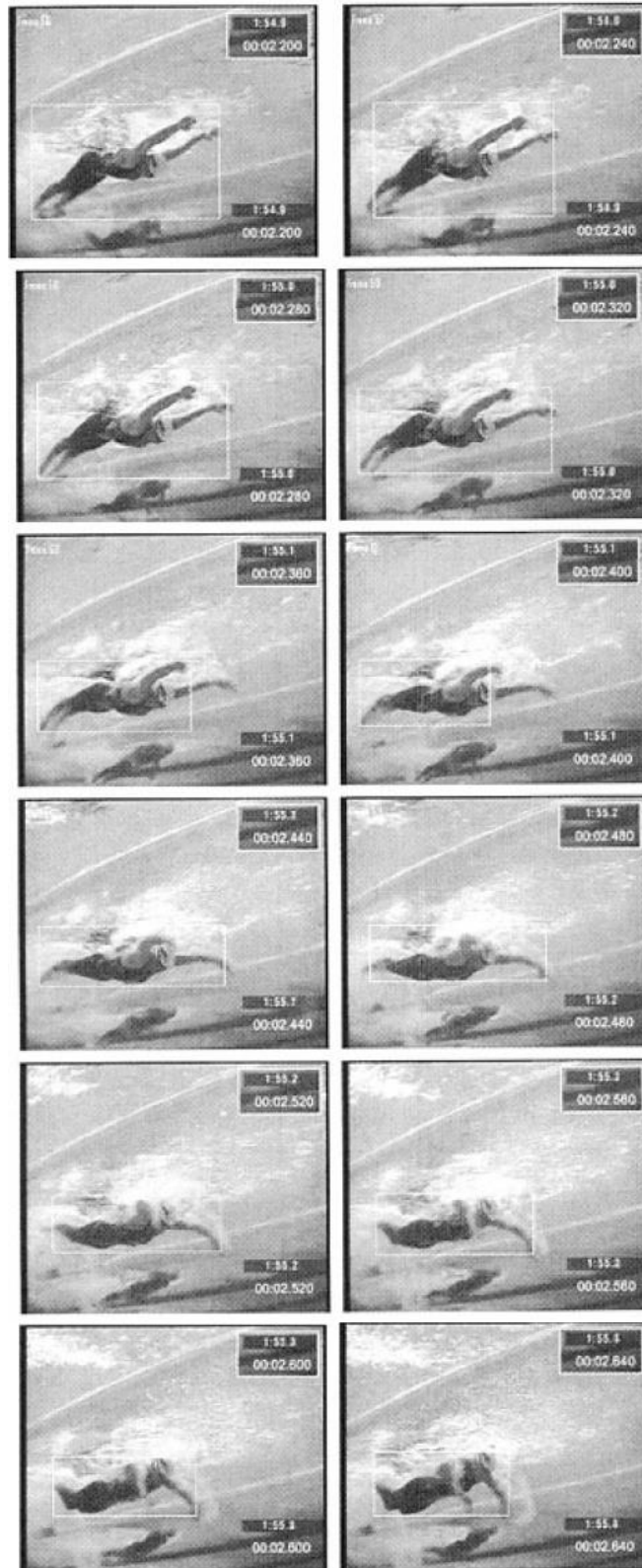


Figure 3. Simulation Results of Prediction and Tracking Based on Kalman Filter

Table 2. Figure 3 the Corresponding External Rectangular Box Information

Frame number	r_0 (Testing)	c_0 (Testing)	h (Testing)	w (Testing)
56 th Frame	102.34	35.89	155.34	210.32
57 th Frame	102.33	45.34	156.34	208.11
58 th Frame	146.34	32.22	110.22	230.32
59 th Frame	35.23	22.12	98.12	241.07
60 th Frame	155.45	24.33	89.01	168.02
61 th Frame	144.65	45.34	76.09	156.12
62 th Frame	145.44	23.78	76.56	178.17
63 th Frame	147.11	23.90	65.05	209.36
64 th Frame	154.89	37.99	67.45	201.56
65 th Frame	154.67	36.07	71.67	189.67
66 th Frame	156.77	40.76	78.34	178.89
67 th Frame	150.66	42.33	76.34	167.09

7. Conclusion

In this paper, the tracking problem of swimming athletes in underwater video system is studied. The method of moving target tracking in video sequences based on Blob analysis is proposed, which is based on Kalman filtering. The Kalman filter uses the feature information extracted from the Blob analysis to initialize the target object and the starting point of the feature matching process. Using the Kalman filter to predict the next frame of the target, the simulation results show that the method is effective.

References

- [1] Z. Zhizhong, "Golf caddie robot visual tracking system research", Chongqing University, (2013).
- [2] Y. Hong, "Research on the technology of human motion tracking of human joints", Harbin Engineering University, (2008).
- [3] L. Tianyu, "Research on the detection and tracking of people and vehicles in the video", Shanghai Jiao Tong University, (2013).
- [4] X. Yu, "A new visual tracking method and its application research", Jiangnan University, (2013).
- [5] Shen, "The bud the national karate team key players technical and tactical tracking analysis and research", Journal of Beijing University of physical education, vol. 9, (2010), pp. 105-109.
- [6] J. Yu, "Research on target detection and tracking technology of sports video based on hybrid algorithm", Soochow University, (2012).
- [7] L. Tao, "Research on the detection and tracking algorithm of ball in soccer video", Xi'an Electronic and Science University, (2012).
- [8] L. Diping, "Method of human body posture detection and motion recognition based on video", Central South University, (2012).
- [9] P. Liang, "Target detection and tracking in soccer video game", Nanjing University of Science and Technology, (2006).
- [10] H. Liqin, "Research on the method of human body's diving motion tracking based on video", Central South University, (2008).
- [11] L. Song, "Research on the technology of character tracking and behavior analysis based on video sequences", Harbin Engineering University, (2010).
- [12] O. Masoud, S. Rogers and N. P. Papanikolopoulos, "Monitoring Weaving Sections", CTS 01-06, ITS Institute, University of Minnesota, (2001).
- [13] N. Funk, "A Study of Kalman Filter applied to Visual Tracking", Project for CMPUT 652, University of Alberta, (2003).
- [14] E. Caevas, D. Zaldivar and P. Gojas, "Kahnan filter for Vision Tracking", Technical Report B 05-12, Freie University Berlin, Germany, (2005).

Author



Ni An, She received her M.S degree from Northeast Normal University. She is a lecturer in Physical Education Department of Northeast Agricultural University. Her research interests include physical education.