15th IAPR International Conference on Machine Vision Applications (MVA)
Nagoya University, Nagoya, Japan, May 8-12, 2017.

Mixture Particle Filter with Block Jump Biomechanics Constraint for Volleyball Players Lower Body Parts Tracking

Fanglu Xie, Xina Cheng and Takeshi Ikenaga
The Graduate School of Information, Production and Systems
Waseda University
Kitakyushu, Japan
waseda-xiefl@fuji.waseda.jp

Abstract

Volleyball player body parts tracking is very important for block or jump height calculation which can be applied to TV contents and tactical analysis. This paper proposes a mixture particle filter with block jump biomechanics constraint based on 3D articulated human model. Using mixture particle filters tracking different body parts can effectively reduce the freedom degree of the human model and make each particle filter track the specific target more accurately. Block jump biomechanics constraint executes adaptive prediction model and likelihood model which can make the particle filter specific for block tracking. The experiments are based on videos of the Final Game of 2014 Japan Inter High School Games of Men's Volleyball in Tokyo. The tracking success rate reached 93.9% for left foot and 93.8% for right foot.

1. Introduction

Nowadays automatic sports analysis becomes an urgent task, due to the popularity of sports and audiences' need for a new sports experiences, which is very important for some developing systems such as highlights extraction system and tactical analysis system. In addition, new sports viewing experiences for audience created through ample TV contents make sports analysis not only an interest research topic but also a valuable commercial opportunity. To meet requirements of TV contents, tracking player body parts and getting the trajectories of hands and feet are prerequisites to calculate the height, speed and other data of players' motions. Our research target is to utilize 3D articulated models to track players' lower body parts.

There have been two conventional methods for tracking players. The first one is tracking the center point of the player. Huang[1] tracks players as points and gets high successful rate but the tracking results are lack of information of body parts. Another one is using articulated model based particle filter to track human body[2][3][4]. However, the state vectors of this conventional work which are of large dimensions lead to inefficient sampling and exponentially increasing computational cost. For tracking lower body parts, del Rincón's algorithm[4] uses 12 parameters as state vectors for each leg. To restrict the human model with high degrees of freedom (DOF), they utilize boundary curve of moving object or silhouette got from foreground segmentation. However, in our research, there are plural number of athletes in volleyball videos. Other moving players' behaviors can also be detected as foreground, which results in huge noises for our research. Without foreground matching, it is very hard to build articulated human models by high DOF. Thus, designing an articulated model building method with fewer DOF is one of the issues we need to solve. In addition, almost all conventional works' experiments are based on simple sequences which include just one person walking or doing other motions without similar objects from other moving people. However, in our research, athletes are doing complex and high-speed block jump within large action range. Furthermore, there are too many other legs of both the teammates and opposite players, which have same features with the targets. So designing a block characteristic based particle filter and reducing the interference of other athletes are very important in our research.

To solve these problems, we propose 3D articulated model based mixture particle filters with biomechanics constraint, which based on a typical particle filter[5]. Mixture particle filter consists of PF1 tracking for hips and PF2 tracking for shanks. Compared with building whole articulated models by a single particle filter[3], mixture particle filter can construct more suitable models for human shape with lower DOF and fewer particles. Block jump biomechanics constraint consists of block characteristics based prediction model of variable step size and biomechanics constraints likelihood model. Both prediction model and likelihood model utilize thigh length, shank length, distance between feet, distance between knees, distances from other players and other distances to limit the 3D articulated models corresponding to real legs of players which implements a specific particle filter tracking lower body parts of block jump and reduces the interference of other players.

The remainder of this paper is organized as follows. Section 2 describes our 3D tracking system. Section 3 and 4 describes details of the tracking method. Experimental results and analysis are given in Section 5, and Section 6 summaries the whole paper.

2. 3D Articulated Model Based Particle Filter

The whole structure of 3D articulated model based particle filter is showed in Fig.1. In the initialization step, we input trajectories of players' torsos got by Huang's method[1] and obtain the 3D positions of hips, knees and feet. In the prediction step, based on biomechanics constraints, we generate particles around feet, knees, and hips in 3D space and combine candidate particles to build lower body models. Then, in the likelihood evaluation step, based on the observation space, we evaluate the possibility of each human model to be the tracking target by 3D distance and HSV color likelihood evaluation. Finally, we select the

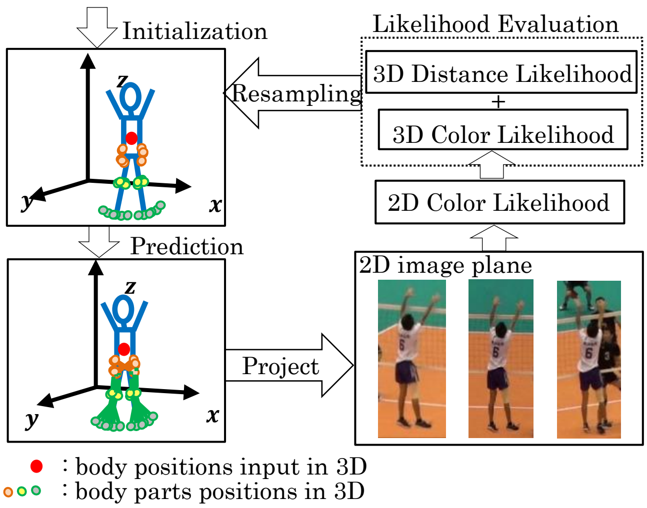

Figure1. Overview of 3D Players Lower Body Parts Tracking

model particle with maximum 3D likelihood as the tracking result. And we do resampling to re-weight particles and iterate the whole process for tracking.

2.1. State Vector and Prediction Model

The designation shows in Fig.2 P, B, H, HM, K, F, L, d, l and r stand for position, body, hip, medium of hips, knee, foot, length, distance, left and right respectively.

State vector X_k at time k contains six 3D position factors. As showed in Fig.2, positions of hips, knees and feet are set as state vectors.

$$X_k = [P_k^{Hl}, P_k^{Hr}, P_k^{Kl}, P_k^{Kr}, P_k^{Fl}, P_k^{Fr}]^T, k \in N \quad (1)$$

In addition, each position factor consists of three dimensional factors in 3D space. And the link of two points in 3D space is called rod model.

$$P_k = (x_k, y_k, z_k), k \in N \quad (2)$$

Torso position P^B acquired by Huang's method is used as input in our framework. At first frame, original positions of hips, knees, feet are got manually in 3D space. Original d^K and d^F are calculated by manually original knees and feet positions. L^{B2H}, L^H, L^T and L^S are fixed lengths calculated by original body parts positions. d^K and d^F are utilized as referential features during the tracking flow.

Then, the prediction model, which indicates how state transits between two consecutive set of particles tacking for human body parts, is formulated as Eq.3

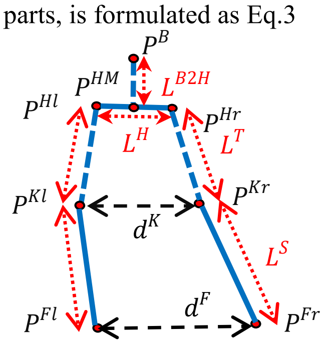

Figure2. 3D Articulated Lower Body Parts Model

$$X_k = F(X_{k-1}) + W_k, k \in N \quad (3)$$

F(X) is a transit function. W_k is the current system noise. The structure of particle filters of this paper is introduced in section 3. The prediction model is introduced in section 4.1.

2.2. Observation Model

Observation model, also called likelihood model, is the criteria to evaluate the similarity between prediction models and real human legs. The observation space I_k is a collection which is constructed by frames at time k by camera m. M is the total number of the cameras.

$$I_k = \{I_k^1, I_k^2, \ldots, I_k^m, \ldots, I_k^M\} \quad (4)$$

In our research, observation model consists of 3D distance likelihood and HSV color likelihood calculated from 2D planes. The whole observation model of likelihood is indicated as below.

$$p(I_k|X_k) = L_{HSV}(X_k|I_k) \times L_{Dis}(X_k|I_k) \quad (5)$$

The observation model of HSV color likelihood is indicated as (6). HSV color likelihood calculation function includes two parts: HSV color histogram based detection method and skin or shoe's color based recognition method. Combing recognition method and detection method improves the robustness of the color evaluation method.

$$L_{HSV}(X_k|I_k) = \prod_{i=1}^{M} L_{HSVD}(X_k; I_k^i) + \prod_{i=1}^{M} L_{HSVR}(X_k; I_k^i) \quad (6)$$

3D distance likelihood model $L_{Dis}(X_k|I_k)$ utilizes length and distance information to evaluate the similarity between predicted human models and real human body in 3D space. This likelihood model is introduced in section 4.2.

3. Mixture Particle Filter

Because of the large DOF of the human model, it is very difficult to set up a prediction model consistent with the real lower body. Using a Gaussian model based single particle filter for tracking requires a large quantity of particles to track the body parts effectively, which results in high computation complexity. In order to build reasonable body models with fewer DOF and particles, this paper proposes a mixture particle filter to track different body parts.

Mixture Particle Filter includes two particle filters, PF1 and PF2. Each particle filter includes its unique state vector, prediction model and observation model. The overall of the mixture particle filter is showed as Fig.3.

PF1 is used for prediction of hips and getting some good prediction results as hips candidates. The state vector of PF1 is defined as (7). For the prediction model, equation (8) shows how the state X_k^H evolves from the previous state. Input data P_k^B is the 3D position of torso at time k. W_k^H is the current system noise for hips prediction. For the observation model, black pants color is used as the color feature and L^{B2H}, L^H are utilized for distance feature.

$$X_k^H = [P_k^{Hl}, P_k^{Hr}]^T, k \in N \quad (7)$$

$$X_k^H = F_H(X_{k-1}^H | P_k^B) + W_k^H, k \in N \quad (8)$$

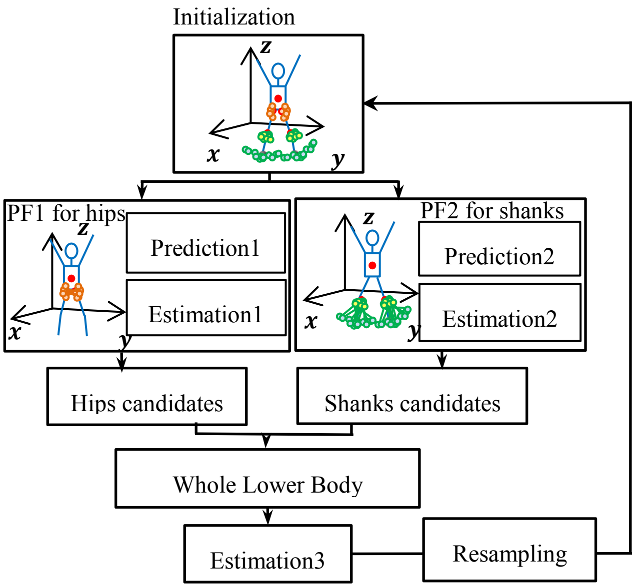

Figure 3. Overview of 3D Players Lower Body Parts Tracking

PF2 is used for prediction of shanks and get shank candidates. The state vector of PF2 is defined as (9). Shanks prediction model is defined as (10). W_k^S is the current system noise for shanks prediction.

$$X_k^S = [P_k^{Kl}, P_k^{Fl}, P_k^{Kr}, P_k^{Fr}]^T, k \in N \tag{9}$$

$$X_k^S = F_S(X_{k-1}^S) + W_k^S, k \in N \tag{10}$$

At last, we estimate whole body models established by the permutation and combination of the hips candidates and shank candidates to get tracking result with high weight.

Conventional works[2][3][4] build total human model for only one time which is difficult to set up models similar to real human shape with few particles. And likelihood values can be easy influenced by each body parts, which results in evaluation accuracy reduction. In our research, the mixture particle filter tracking for each body parts ensures more located rod models on hips or shanks. In addition, during the likelihood evaluation step, unique features for each body part is fully utilized without other body parts' likelihood influence, which helps for higher accuracy.

4. Block Jump Biomechanics Constraint

Compared with particle filter utilizing Gaussian model based prediction model and HSV color likelihood, this paper proposes a method to enhance both the prediction model and the observation model with jump biomechanics constraints.

4.1. Prediction Model

Prediction model showed as (3) is introduced in this section. Eq.(11) shows a transit method utilizing swinging rod model to generate particle P_k around each body parts. This transit method is used in both PF1 and PF2.

$$P_k = f(P_{start}, P_{k-1}, L, u)\,, k \in N \tag{11}$$

We generate particles as the endpoints of swinging rod models from a start point P_{start} in 3D space. L is the 3D length of rod model. u is set as u_- if the tracking target is under the start point, while being set as u_+ when the tracking target is above the start point. Eq.(12) is the expansion equation of (11).

$$\begin{cases} x_k = x_{k-1} \pm \delta_{xy} \cdot \sin\theta_{xy} \\ y_k = y_{k-1} + \delta_{xy} \cdot \cos\theta_{xy} \\ z_k = z_{start} + u * \sqrt{(L \pm \delta_z)^2 - {\delta_{xy}}^2} \end{cases} \tag{12}$$

The ranges of system noise in xy plane and yz plane are adaptive to the block motion. The parameter δ_{xy} is adaptive to θ_{xy}. We divide the transition space in 16 angles at xy plane. The values of δ_{xy} which is used for setting up the largest range of system noise in xy plane are specific by different prediction angles. The transition space is divided by left leg and right leg, based on the positions of torso and hips showed in Fig.4. And the search step length in xy plane is adapted to the distance between knees and feet which avoids overlap of the predicted spaces of the left leg and right leg. And based on several previous torso positions, we judge whether the player is jumping up or falling down. δ_z is adaptive to player's jump condition. We transit more particles above the reference position when player is jumping up but less particles when player is falling down.

For PF1, there are two steps to build the hips models. Step one is generating the particles P_k^{HM} of middle of hips using swinging model introduced as Eq.11 from torso position P^B. The length of this rod model is L^{B2H}. Step two is predicting hips positions P_k^H using swinging rod model based on the center of hip positions we set at step one. The length of this rod model is L^H.

For PF2, there are two steps to build the shanks models. Step one is generating particles P_k^F around feet based on block characteristic using swinging shank rod models from previous knees positions P_{k-1}^K. The length of this rod model is L^S. Step two is predicting the knees positions P_k^K using a reversed swinging shank rod models from the feet positions P_k^F got at step one.

Compared to normal prediction model, fully using the block characteristic in prediction model contributes to building lower articulated body models with high possibility that it can match the real human lower body due to adaptive prediction models based on different angles and motion conditions.

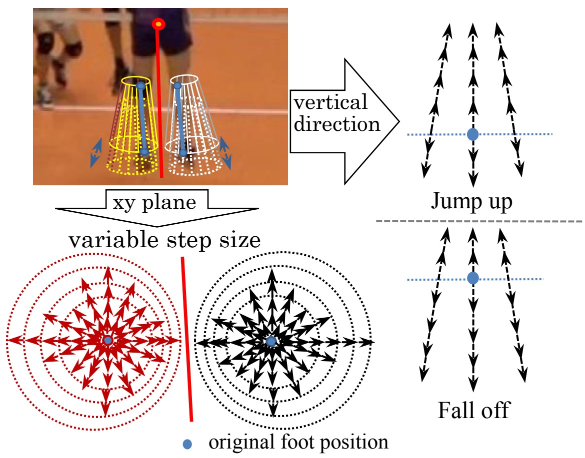

Figure 4. Variable Step-size Prediction Model

4.2. Likelihood Model

In order to get highly accurate tracking results, 3D distance information is a really useful feature for observation model. PF1 utilizes L^{B2H} and L^{H}. PF2 utilizes L^{T}, L^{S}, d^{K} and d^{F}. And both the PF1 and PF2 utilize distance from other players to prevent the tracking results from locating on other players' legs. We also evaluate the distance between current and previous tracking results for each body parts to avoid large error. The total type number is defined as M. All of these distances and lengths are calculated in 3D space I_k. The biomechanics constraint likelihood model is defined as (13)

$$L_{Dis}(X_k|I_k) = \prod_{m=1}^{M} L_{Dism}(X_k|I_k) \qquad (13)$$

Compared to normal likelihood model, implanting 3D distance information to features can effectively prevent the tracking result locating on team mates and tracking result far from the real positions.

5. Experiment result

Volleyball videos used in our experiment are the Final Game of 2014 Japan Inter High School Games of Men's Volleyball in Tokyo Metropolitan Gymnasium. We utilize three cameras located at the left corner, right corner and middle behind the target team to realize the requirement of 3D tracking. We delete some sequences that target players are totally occluded by other players which are impossible to track the target. The total video sequences number is 22. Plural players are blocking at the same time in 5 sequences while other sequences consisting of one athlete. And total sequences contains 1798 frames in all. One sequence stands for one flow of block motion. The final resolution of videos is 1920×1080 and the frame rate is 60 frames per second. We set the shutter speed as 1000/s to avoid the motion blurs. Our proposal is implemented using C++ and OpenCV 2.4.10.

In order to show the contribution of each proposal, we evaluate proposals step by step. First, we evaluate the framework, using Gaussian model based single particle filter with HSV color likelihood model. Then we evaluate the experimental results with mixture particle filter. At last, we evaluate the experimental results combining mixture particle filter with block jump biomechanics constraint.

To make the tracking result be observed visually, we draw the projected 2D tracking results using four green rectangles to cover thighs and shanks in each directional video sequence showed as Fig.5. We define the evaluation methods divided by 6 points, positions of hips, knees and feet. If the side of projected rectangle exactly covers or covers part of the tracking target within a certain error range, this frame would be set as a successful frame.

$$\text{success rate} = \frac{\sum \text{success frame}}{\sum \text{frame}} \times 100\% \qquad (14)$$

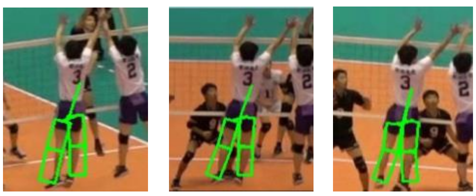

Figure 5. Projected 2D Tracking Windows

Table I. Evaluation result

Success Rate (%)	Hips		Knees		Feet	
	left	right	left	right	left	right
Framework[1]	84.8	82.6	66.3	65.9	67.1	66.9
P1[2]	94.7	95.1	81.7	76.8	78.4	80.6
P1&P2[3]	97.7	98.5	91.4	92.2	93.9	93.8

[1]Framework: Gaussian model based single particle filter;
[2]P1: Gaussian model based mixture particle filter;
[3]P2: Block jump biomechanics constraint.

The method described in this paper can't achieve perfect tracking accuracy because of several factors. Opposite players' legs and large motion range influence the tracking system a lot. The player's shadow, having the color similar to skin color. For future, after solving these problems, we plan to transit this framework to track upper body parts.

6. Conclusion

In this paper, aiming at achieving high success rate of tracking player feet in 3D space, we propose a Mixture Particle Filter with Biomechanics Constraint to realize fitted human lower body parts tracking. Mixture particle filters are used for tracking different targets respectively which can reduce the degrees of freedom and build more accurate rod models located on each body parts with fewer particles. Biomechanics Constraint for block jump enhances both the prediction model and likelihood model which contributes to building lower articulated body models with high possibility identical with the real human lower body. The tracking success rate reached 93.9% for left foot and 93.8% for right foot. For future work, the current tracking system can be applied to tracking of hands or other actions such as spike, and serve, which is very useful for enriching the TV contents.

References

[1] S. Huang, X. Zhuang, N. Ikoma and T. Ikenaga: "Particle filter with least square fitting prediction and spatial relationship based multi-view elimination for 3D Volleyball players tracking," *2016 IEEE 12th International Colloquium on Signal Processing*, pp. 28-31, 2016.

[2] M.W. Lee, I. Cohen, and S.K. Jung: "Particle filter with analytical inference for human body tracking," *Proceedings Workshop on Motion and Video Computing*, pp. 159-165, 2002.

[3] S. Kim, C.B. Park, S.W. Lee:"Tracking 3D human body using particle filter in moving monocular camera," *18th International Conference on Pattern Recognition (ICPR'06)*, vol.4, pp.805-808, 2006.

[4] J.M. del Rincón, D. Makris, C. O. Uruñuela, J. C. Nebel: "Tracking human position and lower body parts using Kalman and particle filters constrained by human biomechanics," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol.41, no.1, pp.26-37, 2011.

[5] M.S. Arulampalam, S. Maskell, N. Gordon and T. Clapp: "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174-188, 2002.