5-2012

A Novel and Effective Short Track Speed Skating Tracking System

Yuxuan Wang
Utah State University

A NOVEL AND EFFECTIVE SHORT TRACK SPEED SKATING TRACKING SYSTEM

by

Yuxuan Wang

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Computer Science

Approved:

_____	_____
Dr. Heng-Da Cheng Major Professor	Dr. Brett E. Shelton Committee Member
_____	_____
Dr. Stephen J. Allan Committee Member	Dr. Curtis Dyreson Committee Member
_____	_____
Dr. YangQuan Chen Committee Member	Dr. Mark R. McLellan Vice President for Research and Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2012

ABSTRACT

A Novel and Effective Short Track Speed Skating Tracking System

by

Yuxuan Wang, Doctor of Philosophy

Utah State University, 2011

Major Professor: Dr. Heng-Da Cheng
Department: Computer Science

This dissertation proposes a novel and effective system for tracking high-speed skaters. A novel registration method is employed to automatically discover key frames to build the panorama. Then, the homography between a frame and the real world rink can be generated accordingly. Aimed at several challenging tracking problems of short track skating, a novel multiple-objects tracking approach is proposed which includes: Gaussian mixture models (GMMs), evolving templates, constrained dynamical model, fuzzy model, multiple templates initialization, and evolution. The outputs of the system include spatial-temporal trajectories, velocity analysis, and 2D reconstruction animations. The tracking accuracy is about 10 cm (2 pixels). Such information is invaluable for sports experts. Experimental results demonstrate the effectiveness and robustness of the proposed system.

(68 pages)

PUBLIC ABSTRACT

A Novel and Effective Short Track Speed Skating Tracking System

Short track speed skating lends itself to intense competitions with a strong visual impact. Thus, the sport has become increasingly popular. In fact, in 1992, short track speed skating became an official Winter Olympic sport with four events, and four more events were added in 2002. Because of the sport's growing popularity, there is a high demand from both coaches and TV broadcasters for a means of automatically gathering competition data such as trajectories, velocities, and 2D reconstruction animation. We call this vision-based sports video analysis.

In competitive short track speed skating, multiple skaters skate together on an ice track the size of a hockey rink. Different from traditional longer track speed skating, during short track competition, there are no visible lanes, and skaters usually cluster in groups, especially around curves. Due to this crowding, skaters' velocity can change dramatically although still moving at a high rate of speed. Additionally, skaters on the same team wear identical uniforms, and the uniforms of different teams can be nearly identical. The clustering of the skaters while traveling at high speeds, and little visual difference between skaters can make sports video analysis difficult to achieve.

While there are many sports video analysis methods currently available, none of them lend themselves well to short track speed skating. This dissertation addresses this problem.

Yuxuan Wang

This work is dedicated to

my parents, Weipu Wang and Zhaoqin Gong,

my wife, Juan Shan,

and my son, Kevin Wang.

ACKNOWLEDGMENTS

First of all, I would like to express my sincere and heartfelt gratitude to my advisor, Dr. Heng-Da Cheng. Without his consistent and professional academic guidance, I would not have been able to finish this dissertation. His suggestions illuminated my research direction, and his encouragement helped me to overcome obstacles. I am also very appreciative of my committee members, Dr. Yangquan Chen, Dr. Stephen Allan, Dr. Brett Shelton, and Dr. Curtis Dyreson, for their comments, advice, and contributions to my research.

I wish to thank my classmates and friends who assisted me during my research at Utah State University. Thanks go to the group members of the CVPRIP lab, Wen Ju, Yanhui Guo, Ming Zhang, and Chenguang Liu, for their help and cooperation. Special thanks go to my neighbor Yanxu Liu, who provided much needed help to my life when I was writing the final draft of my dissertation.

Particularly, I would like to thank my wife, Juan Shan, a previous member of the CVPRIP group. She gave me lots of useful suggestions, moral support, and a most beautiful two-year-old son. My sincere appreciation goes to my parents and my parents-in-law, who helped us by taking care of our son, Kevin Wang. Their help enabled me to focus on my dissertation research. Thanks go to my dear little son, Kevin, who is the most beloved person in my family and also gives me an endless source of power and encouragement.

Yuxuan Wang

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Short Track Speed Skating

Short track speed skating is a form of competitive ice speed skating. Multiple skaters, usually four to six, skate together on an ice track with a circumference of 111.12m. The size of the rink itself is the same as a hockey rink -- $60m \times 30m$.

Different from traditional longer track speed skating, during the competition no visible lanes are employed in short track speed skating. Skaters usually cluster in groups especially around curves. Due to the crowding, the skater's velocity can change dramatically although still moving at a high rate of speed. For example, in a sprint, a professional male skater can reach about 60km/h, and the typical average velocity of a professional female skater is about 36km/h, in a 500m race.

Smaller ice tracks lend themselves to intense competitions with a strong visual impact. Consequently, short track speed skating has become more and more popular in many countries. In 1992, short track speed skating became a regular Winter Olympic sport. And the program was expanded from four short track events in 1992 to eight in 2002.

Both individual and team tactics are very important in short track speed skating competitions. Usually, a smarter skater can defeat a faster skater through better tactics. For example, having a teammate on the rink provides a huge advantage over others who do not.

There is a high demand from both coaches and TV broadcasters to automatically gather competition data such as trajectories, velocities, and 2D reconstruction animation.

Such data not only enhances TV coverage of competitions, but it also helps coaches to devise tactics, monitor training processes, and evaluate statistics.

1.2 Vision-based Sports Video Analysis

Vision-based sports video analysis has many applications including velocity analyzing, 2D/3D-reconstruction, special events finding, tactics analyzing, automatic highlight identification, video annotation, graphical object overlaying, etc. Some of these applications improve user experience, while others provide invaluable statistical and visualized information to athletes and coaches.

1.2.1 Review of Current Approaches

Following are brief reviews of some current video analysis approaches. Soccer, the most popular worldwide team sport, has attracted a lot of video analysis research, and many applications have been developed. In [1], soccer players are tracked by different features such as texture, color statistics, movement vectors, etc. In [2], color image segmentation is applied to extract meaningful regions to represent players. The method proposed in [3] tracks multiple players obtained from a single fixed camera, and a Kalman filter is employed to improve the prediction of player movement. The authors of [4] proposed an automatic multiple-player detection system. Background subtraction, a Haar feature, and Markov chain Monte Carlo (MCMC) data association were used. However, as described in D'Orazio and Leo's soccer video analysis survey paper [5], developing more robust algorithms for player detection is still a challenging problem, especially, when players are occluded with each other in a long time period. Furthermore, some of these methods require a camera with no abrupt motion.

Many computer vision systems have been also developed for analyzing other sports videos. In [6], tracking techniques were employed to enhance the broadcast of tennis matches. Recently, [7] proposed a data association algorithm to track tennis balls by mounting cameras directly above the court. Color features and a template are combined to track the players. Other sports, such as squash [8], handball [9], and basketball [10], use the same technique. However, the method suffers from two problems: first, the cameras must be placed above the playing court. Second, the occlusion problem is not solved effectively. The authors of [11] developed an automatic annotation system for American football. The "close-world" is defined as a spatial-temporal region in which the context is sufficient for detecting all objects in that region. Motion blobs are used to track players by "close-world" analysis. However, the approach cannot deal with group players as individuals; instead, players are considered as a single unit. In [12], a color-based sequential Monte Carlo tracker is employed to track hockey players. It is a single player tracking system, and its accuracy is only 0.3m~1m. When occlusion happens, the tracker can lose the target. In [13], the authors proposed a unified dual-mode two-way Bayesian inference approach. The method is applied to soccer, basketball, and hockey. Partial occlusion is handled by forward filtering and backward smoothing. However, the players with similar appearance or a long time occlusion might lead to tracker failure.

Besides sports video analysis, there are many other tracking applications including video surveillance [14], smart environments [15, 16], pedestrian tracking [17, 18], face tracking [19, 20], etc. In addition, mean shift algorithm [21, 22] is a simple and robust method by finding local minima of a similarity measure between kernel density

estimations. A color distribution-based particle filter has good performance on tracking non-rigid objects with non-linear and non-Gaussian motion [23, 24]. However, all of the above approaches are not capable of dealing with complicated conditions, such as amorphous, erratic motion, fast speed, similar appearance, long time occlusion in a large-scale, complex and dynamic scene, etc., of short track skating

1.2.2 Challenging Issues

Compared with the above approaches, tracking short track speed skaters by video encounters several open-ended, challenging problems:

1) Skaters move rapidly and change directions frequently;

2) The camera's motion is not smooth;

3) The size of a skater in the scene varies dramatically;

4) The shape of a skater deforms frame by frame;

5) Long time occlusion and complex group situations, especially on the curves of the rink, are inevitable;

6) Skaters on the same team wear identical uniforms;

7) The speeds among skaters are similar;

8) There is no detectable intersection point between field lines. Thus, no wire frame model such as those utilized in soccer or football video analysis, can be employed to register the short track rink.

1.2.3 Previous Work

We are the first group to solve these challenging problems, and did some preliminary work [25]. Based on RANSAC [26, 27] algorithm, a rink registration method

was developed and a histogram-based hierarchical model was proposed to track the skaters. The system successfully tracked a single player in a 500m match. However, some challenging problems remained unsolved:

1) It was only a single-person tracking system;

2) It was difficult to find key frames for calculating homography automatically;

3) Long time occlusion among skaters, especially when the occluded skaters wore the same kinds of uniforms, caused missed tracks;

4) The system relied on helmet detection; however, the method sometimes failed due to the small size of the helmet or other occlusions. In addition, skaters sometimes wore helmets with the same color as their clothes, thus making distinguishing the helmet difficult.

1.3 The Proposed Method

In this paper, we propose a novel and effective short track speed skating tracking system (STSSTS, or STS2). A single panning camera is used to obtain videos, as in [25]. The camera is mounted at the audience section of the stadium as close as possible to the center in order to reduce projection error. No camera parameters, height, or angle are required. Zooming is abandoned due to little texture information on the rink, which makes both recording and registration much more difficult.

Such a system is easily deployed by coaches themselves. Its low cost is another advantage. Finally and most importantly, the system answers all the challenges listed in Section 1.2.2.

The novelties and contributions of this work are:

1) Based on the characteristics of a panning camera, a new technique is developed to automatically discover key frames for building the panorama. These key frames are almost equally distributed along a horizontal direction. For each frame, the algorithm finds a proper key frame to calculate the homography automatically.

2) By using GMMs combined with an evolving template to model objects (skaters, referees, and rink), fuzzy memberships of the pixels are calculated according to corresponding objects. Memberships are iteratively updated to find the optimum position of each object.

3) The position found by memberships is considered as the initial position of the template for each object. Then, a multiple-templates evolution algorithm is developed. By doing this, even occluded players from the same team can be tracked properly.

4) Based on the rink parameters, a constrained dynamic model is proposed to handle dynamic issues.

5) The proposed algorithm can handle situations like dramatic changes in the size or shape of the skaters.

6) By using homography between two adjacent frames, a region of interest (ROI) that only contains skaters is generated automatically.

7) STS2 can automatically output lots of invaluable data like velocities, trajectories, and 2D reconstruction animations.

Both 500m men's and a 500m women's world cup short track videos are employed for testing. Both videos have multiple skaters from the same team. 2D

reconstruction animations, trajectory detection, and velocity curve analysis demonstrate the robustness and effectiveness of the proposed system.

The proposed system, STS2, consists of two sub-systems: a registration system and a tracking system. Chapter 2 describes registration system, and Chapter 3 describes the tracking system. Experimental results and discussions are given in Chapter 4, and conclusions are provided in Chapter 5.

CHAPTER 2

REGISTRATION SYSTEM

Velocity analysis and 2D reconstruction require real world positions for skaters in each frame. The task of the registration system is to map the frames in a video to a pre-defined real world coordinate system.

2.1 Related Work

In [26-28], similar requirements are needed for soccer and tennis video analyses. Registration is done by detecting lines and intersection points. In other application such as American football video analysis [11], more contextual information such as yard numbers and logos are also employed. However, no obvious lines with intersection points can be detected on the short track rink. Five starting lines and one finishing line are all parallel with each other. Fourteen markers are placed on two curves, and they are often occluded by skaters. These lines and markers stand little chance of appearign in the same frame. Thus, a simple wire/point model is not capable of registering a short track speed skating rink.

In [12], based on Kanade-Lucas-Tomasi (KLT) [29, 30] features, RANSAC algorithm is employed to calculate the frame-to-frame homographies. However, projection errors are accumulated over time, which cause a large projective distortion for long time tracking.

By finding key frames based on the frame matching algorithm [31, 32], accumulative error can be avoided [25]. However, matching n frames has complexity

$O(n^2)$ or $O(nlogn)$ with some constraints [31, 32]. In this dissertation, we propose a new method to reduce such complexity to $O(n)$.

2.2 Homography

A homography H is a projective transformation matrix with size 3×3. H is an invertible transformation, mapping an image to another. Corresponding points x from image 1 and x' from image 2 hold the relation $x' = Hx$. Because the homography H has eight degrees of freedom (DOF), four point correspondences can determine H. However, more pairs of points should be employed for the real world task, since no perfect correspondences can be detected. Calculating homography between two frames is the basic component of registration.

Harris corner detector [33] and KLT are two popular feature extraction methods for points. They are invariant to both rotation and translation, but not invariant to affine or projective transformations [34]. They are not suitable for our system as indicated in [25], since there are other kinds of transformations.

In this paper, we use scale invariant feature transform (SIFT) [35] to find interest points. SIFT features are invariant to scale and rotation, and robust to illumination change. They have been successfully applied to object recognition [36] and panorama recognizing [37].

As in [25], a simple putative matching method is employed to obtain primitive pairs of the interest points detected by SIFT features between two frames. RANSAC

algorithm [31, 32] can discard false matched pairs (outliers). Finally, homography H can be calculated based on the remained pairs (inliers).

2.3 Proposed Registration Method

Registration Algorithm. Input: Short track speed skating video with n frames.

Step 1: Extract SIFT features for all frames, and denote s_i as the set of interest points detected from the ith frame.

Step 2: Find putative matched pairs $p_{i-1,i}$ based on s_{i-1} and s_i.

Step 3: Define the homography between two adjacent frames $i-1$ and i as $H_{i-1,i}$. Calculate $H_{i-1,i}$ based on $p_{i-1,i}$ using RANSAC algorithm.

Step 4: Denote $c=(x_c, y_c, 1)'$ as the homogeneous coordinates of the center point of the frame. Denote $c_i=(x_i, y_i, 1)'$ as the transformed point from the center point of frame $i-1$. Then, c_i can be calculated by equation:

$$c_i = \aleph \times (H_{i-1,i} \times c) \tag{1}$$

where $\aleph$ is a scalar to make sure that the 3rd term of c_i is equal to 1.

Step 5: Calculate the horizontal offset of the center point in frame i by:

$$x_i^{offset} = x_i - x_c \tag{2}$$

Step 6: Calculate the accumulative horizontal offset for frame i by:

$$\begin{cases} acc_i = acc_{i-1} + x_i^{offset}, & \text{if } i > 1 \\ acc_i = 0, & \text{if } i = 1 \end{cases} \tag{3}$$

Step 7: Assume the number of key frames is K_n. Then, the interval between two neighbor key frames is:

$$Int = (\max(acc) - min(acc)) / (K_n + 1) \tag{4}$$

Denote $acck_j$ as the optimum accumulative horizontal offset for the jth key frame:

$$acck_j = \min(acc) + Int * j \tag{5}$$

Step 8: Find the frame index Kf_j for the jth key frame:

$$Kf_j = \underset{i=1,,n}{argmin} | acc_i - acck_j | \tag{6}$$

Step 9: Find the suitable key frame index K_i for the ith frame:

$$K_i = \underset{j=1,..,K_n}{argmin} | acc_i - acck_j | \tag{7}$$

Step 10: Use all key frames to build the panorama according to [35] and obtain the homography $H_{Kf_j,pa}$ transforming the jth key frame to the panorama.

Step 11: Map the panorama to the real world coordinate system of the rink, and obtain the homography $H_{pa,rink}$.

Step 12: Find putative matches $p_{i,Kf_{K_i}}$ based on s_i and $s_{Kf_{K_i}}$. Compute homography H_{i,K_i} transforming the ith frame to its corresponding key frame using RANSAC.

Step 13: Compute homography H_i mapping the ith frame to the real world coordinate system of the rink by:

$$H_i = H_{pa,rink} \times H_{Kf_{K_i},pa} \times H_{i,K_i} \tag{8}$$

The homography H_i is needed to register the ith frame. The characteristics of the panning camera are utilized to simplify the problem. Since the camera is panning horizontally from left to right or from right to left without zooming, the accumulative horizontal offset (acc) is introduced by measuring the shift of the center points between two adjacent frames. Figure 2.1 shows acc from a 500m short track speed skating video, $K_n = 15$. From Step 4 to Step 8, 15 key frames are obtained and displayed with green circles. Dash lines represent the optimum accumulated horizontal offset $acck_j$ for

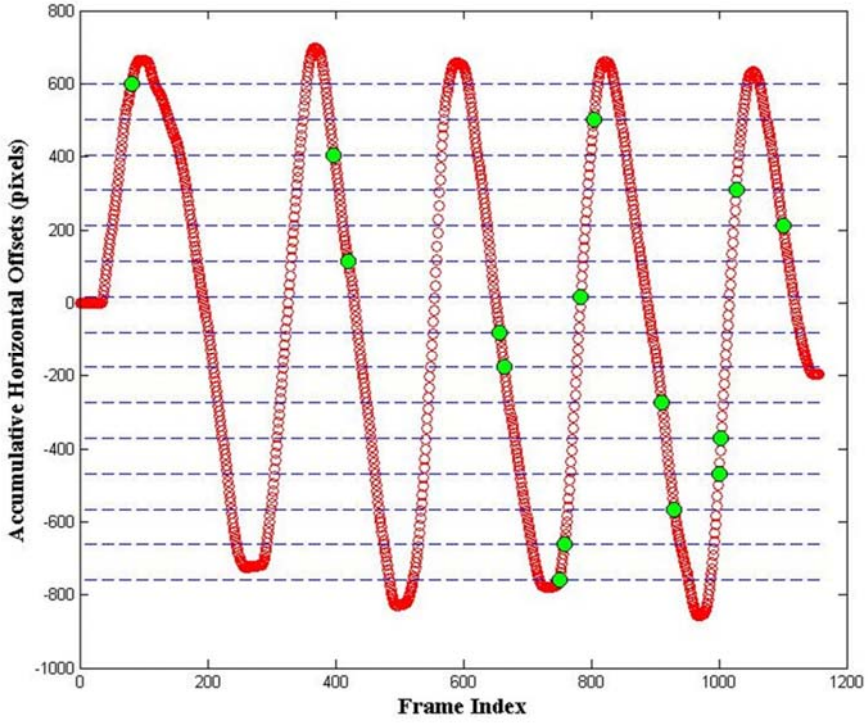

Figure 2.1. Determining the indexes of the key frames based on accumulative horizontal offsets.

each key frame j. The green circles are the nearest frames to the dash lines.

The advantages of the proposed registration system are:

1) Key frames are equally distributed along the horizontal direction.

2) In the proposed registration system, Steps 2 and 3 need $n-1$ frame matches to obtain the homography $H_{i-1,i}$. Step 10 needs $15log(15)$ frame matches to build the panorama based on 15 key frames. Steps 12-13 need n frame matches to obtain the homography H_{i,K_i}. The proposed registration system needs $(n-1)+15log(15)+n$ matches which is $O(n)$. Compared with [25], the proposed system is much more efficient.

3) The idea of equally dividing the interval between the global maximum and minimum of the accumulative horizontal offsets can be employed for many other applications, such as soccer, American football and hockey, with different kinds of motions for the panning cameras.

Figure 2.2 shows the flowchart of the algorithm, using a 500m short track speed skating video. Notice that automatically detected key frames cover the whole scenario, and the panorama is built successfully. In Step 10, the 14 markers are manually picked out in order to map the panorama to the coordinate system of the rink. In Step 11, the range of 2D coordinate system is 1200×600. The size of the rink is $60m\times30m$ and the locations of the markers can be calculated based on [38]. Thus, one pixel in the coordinate system represents 5cm in the real world.

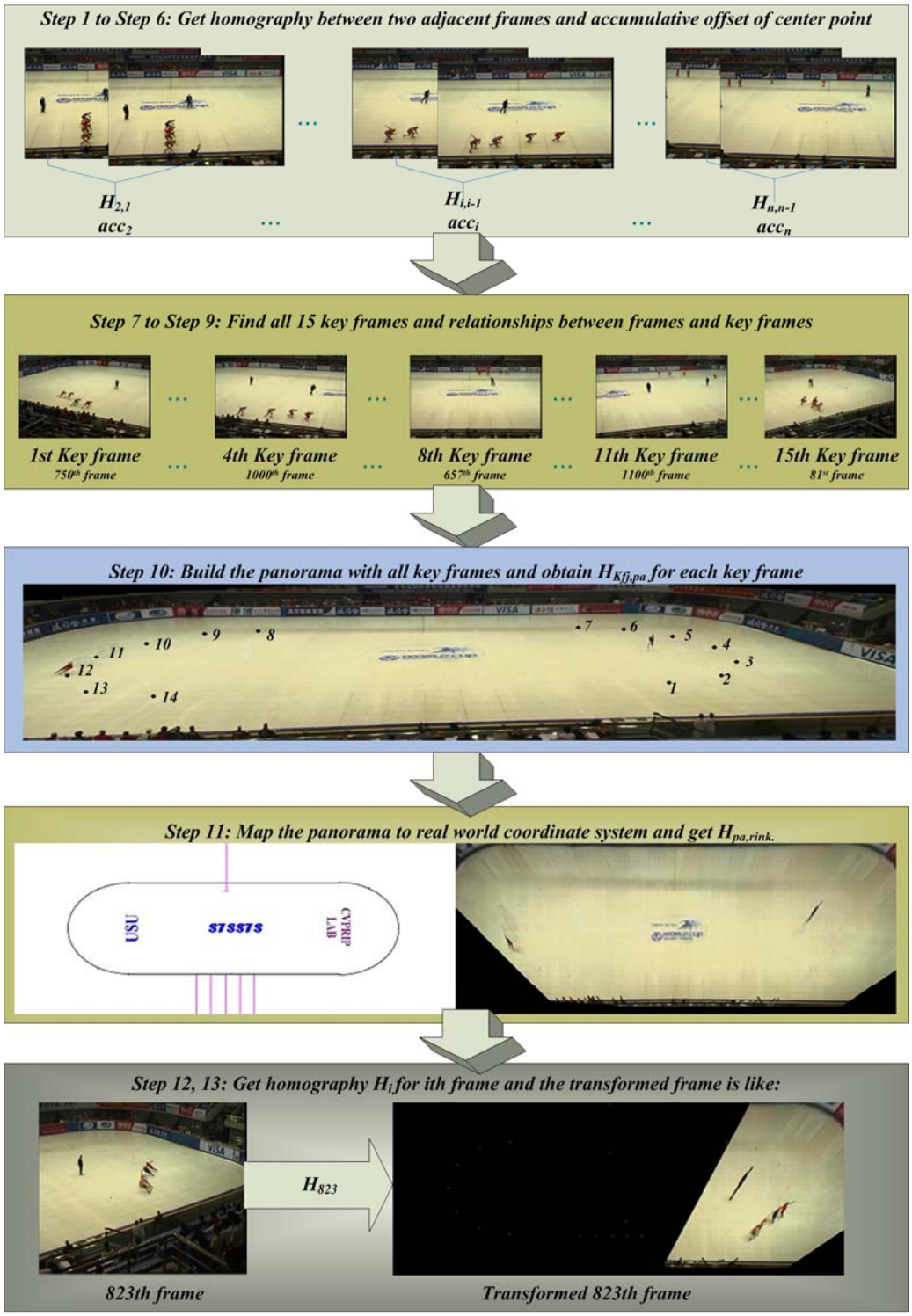

Figure 2.2. The flowchart of the proposed registration algorithm.

CHAPTER 3

TRACKING SYSTEM

3.1 Appearance Model

The uniform of a skater may have several colors. To get around this, we manually outline the silhouettes of all skaters, as well as the referee appearing in the first frame, as shown in Figure 3.1. The helmet of the skater is ignored. Thus, the proposed algorithm has no hierarchical constrains like in [25], and can be employed for more generalized applications.

In Figure 3.1, the first and second skaters wear identical uniforms of red and black. The uniform of the third skater has three colors, namely, blue, black, and yellow. The fourth skater has pink and black. These colors are, of course, not pure colors. However, based on subjective observation, each uniform can be considered as consisting of certain number of color components. Each color component is assumed to be normally distributed.

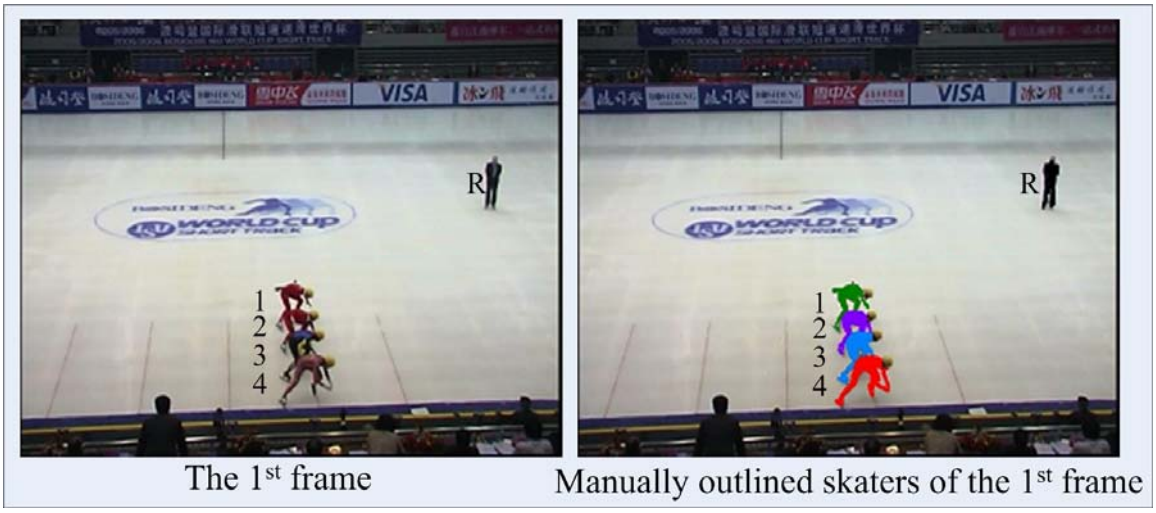

Figure 3.1. Initialization of the first frame.

Color-based GMMs was employed in many computer vision applications like background modeling [39] and shadow detecting [40]. In this dissertation, we use Gaussian mixture models (GMMs) to model the relationship between colors and objects. The objects can be the skaters, referees, or even the rink. The number of mixture Gaussians is determined by the number of color components.

Define p_{ob_j} as the Gaussian mixture probability density function of object j, where $j \in \{\text{SK, RF, } rink\}$, where $\text{SK} = \{1,2,3,...\}$ represents the set of skater indexes; $\text{RF} = \{R_1, R_2, ...\}$ represents the set of referee indexes; and *rink* is the index of the rink. Denote I^i as the *i*th frame and $v^i(x,y) = (r,g,b)'$ as the transpose of the RGB values of $I^i(x,y)$. Then:

$$p_{ob_j}\left(v^i(x,y)\right) = \sum_{k=1}^{m_j} c_{j,k} G[\mu_{j,k}, \Sigma_{j,k}, v^i(x,y)] \tag{9}$$

$$G\left[\mu_{j,k}, \Sigma_{j,k}, v^i(x,y)\right] = \frac{1}{2\pi |\Sigma_{j,k}|^{1/2}} e^{-1/2\left(v^i(x,y)-\mu_{j,k}\right)' \Sigma_{j,k}^{-1}\left(v^i(x,y)-\mu_{j,k}\right)} \tag{10}$$

where m_j is the number of mixture Gaussians for object j; $c_{j,k}$, $\mu_{j,k}$ and $\Sigma_{j,k}$ indicate the mixture rate, mean vector, and covariance matrix of the *kth* Gaussian for object j, respectively.

Define $\Omega_{ob_j}^i$ as the evolving template of object j in frame i. Note that $\Omega_{ob_j}^1$ is the manually outlined template in the first frame. We have:

$$\begin{cases} \Omega_{ob_j}^{i}(x,y)=1, \text{ if } I^{i}(x,y)\in ob_j \\ \Omega_{ob_j}^{i}(x,y)=0, \text{ if } I^{i}(x,y)\notin ob_j \end{cases} \tag{11}$$

Denote $\mathfrak{U}$ as the uniform index. As in Figure 3.1, there are three kinds of skater uniforms $\mathfrak{U}_1, \mathfrak{U}_2$,and $\mathfrak{U}_3$. Define U_{ob_j} as the uniform worn by object *j*. We have $U_{ob_1}=U_{ob_2}=\mathfrak{U}_1$, $U_{ob_3}=\mathfrak{U}_2$, and $U_{ob_4}=\mathfrak{U}_3$. All referees share the same kind of uniforms marking as $U_{ob_{R_j}}=\mathfrak{U}_R$. The rink is also considered as a kind of uniform and denoted by $U_{ob_{rink}}=\mathfrak{U}_{rink}$. Let uc_k represent the number of color components for $\mathfrak{U}_k$. Note that the rink template $\Omega_{ob_{rink}}^{1}(x,y)$ covers both the audience section and the rink. Let $uc_{rink}=2$.

K-means algorithm [41] and maximum likelihood estimation (MLE) are employed to calculate the parameters of GMMs.

For a given uniform $\mathfrak{U}_k$, pixels belonging to this uniform can be generated by:

$$\Omega_{u_k}=\bigcup_{U_{ob_j}=\mathfrak{U}_k}\Omega_{ob_j}^{1} \tag{12}$$

Apply K-means algorithm to $v^{1}(\Omega_{u_k})$ with $K=uc_k$. Assume that we obtain uc_k output clusters denoted as $\Omega_{u_k}^{l}$, and the number of elements for each cluster denoted as $n_{\mathfrak{U}_k}^{l}$, where $l=1,2,...,uc_k$. For ob_j with $U_{ob_j}=\mathfrak{U}_k$ and $k\neq rink$, parameters for the *lth* Gaussian can be calculated based on maximum likelihood estimation (MLE):

$$m_j=uc_k \tag{13}$$

$$c_{j,l} = \frac{n_{\Upsilon_k}^{l}}{\sum_{q=1}^{uc_k} n_{\Upsilon_k}^{q}} \tag{14}$$

$$\mu_{j,l} = \frac{1}{n_{\Upsilon_k}^{l}} \sum_{\Omega_{\Upsilon_k}(x,y)=1} v^1(x,y) \tag{15}$$

$$\Sigma_{j,l} = \frac{1}{n_{\Upsilon_k}^{l}} \sum_{\Omega_{\Upsilon_k}(x,y)=1} \left(v^1(x,y) - \mu_{j,l}\right)\left(v^1(x,y) - \mu_{j,l}\right)' \tag{16}$$

Based on the observation that the rink covers the largest area of the frame, if $k = rink$, only the cluster with the maximum number of elements is kept for calculating GMMs. To simplify the task, the audience section is removed in the ROI-generation step for every frame.

Finally, a *log* transformation is utilized, and *eps* is added on p_{ob_j} to avoid the underflow problem. Transformed probability is denoted by w_{ob_j}:

$$w_{ob_j}(v^i(x,y)) = -log(p_{ob_j}(v^i(x,y)) + eps) \tag{17}$$

where *eps* is the minimum value of the machine and w_{ob_j} is considered as the weighted GMMs template.

3.2 ROI Generation

For the current input frame i, ROIs are automatically generated for further processing. These ROIs include one region of all skaters and several regions for different referees. An ROI is represented by an axis-aligned minimum rectangular bounding region.

Define $\mathcal{T}(I,H)$ as the transforming function, which transforms image I based on homography H. We can then obtain a transformed previous frame I_t^{i-1} by $\mathcal{T}$ and the adjacent homography $H_{i-1,i}$ which is defined in Step 3 of Section 2.3:

$$I_t^{i-1} = \mathcal{T}(I^{i-1}, H_{i-1,i}) \tag{18}$$

Define R^i as the rink template for the *ith* frame:

$$\begin{cases} R^i(x,y) = 1, & \text{if } I^i(x,y) \in \widetilde{AS} \\ R^i(x,y) = 0, & \text{if } I^i(x,y) \in AS \end{cases} \tag{19}$$

where AS represents the region of audience section and $\widetilde{AS}$ represents the rink region. To obtain R^i, we employ the homography H_i which is obtained in Step 13 of Section 2. The homography H_i transforms the *ith* frame to the real coordinate system of the rink, which only includes the rink region. The following two equations are utilized to calculate R^i.

$$(x', y', 1)' = \aleph \times (H_i \times (x, y, 1)') \tag{20}$$

$$\begin{cases} R^i(x,y) = 1, & \text{if } (x', y') \text{ is inside the coordinate system of the rink} \\ R^i(x,y) = 0, & \text{if } (x', y') \text{ is outside the coordinate system of the rink} \end{cases} \tag{21}$$

Next, a simple frame differencing algorithm is executed between I_t^{i-1} and I^i based on Euclidian distance among RGB values with a threshold setting of 50, and a motion template D^i is obtained. Remove the region of audience section in D^i:

$$D^i = D^i \oplus (D^i \bigcap R^i) \tag{22}$$

where $\oplus$ represents the XOR operator.

As an example, frame 267 and 268 are shown in Fig 3.2A) and B). The motion template D^{268} is shown in Fig 3.2C). Note that D^i ($i = 268$) records the motion information between frame i and frame $i-1$. However, motion alone cannot determine the ROI. Define a transformed template $\Lambda^i_{ob_j}$, which is transformed from $\Omega^{i-1}_{ob_j}$ in frame $i-1$. We have:

$$\Lambda^i_{ob_j} = \mathcal{T}(\Omega^{i-1}_{ob_j}, H_{i-1,i}) \tag{23}$$

For example, the transformed templates for all objects are displayed in Fig 3.2D).

Define a union mask $\Psi^i = \Lambda^i_{ob_j} \bigcup D^i$. Next, employ a 5×5 dilating operation on Ψ^i as shown in Fig 3.2E). Denote ψ^i_l as the *lth* connected component of Ψ^i. Denote $\Lambda^i_{ob_{skaters}}$ as the union of the templates for all skaters. If there is a transformed referee template $\Lambda^i_{ob_{R_k}}$ intersected by any skater mask. , then $\Lambda^i_{ob_{skaters}} = \Lambda^i_{ob_{R_k}} \bigcup \Lambda^i_{ob_{skaters}}$. All objects included by $\Lambda^i_{ob_{skaters}}$ are denoted as $o^i_{skaters} = \{ob_j \mid \Lambda^i_{ob_j} \subset \Lambda^i_{ob_{skaters}}\}$. Note that if $\Lambda^i_{ob_{R_k}} \subset \Lambda^i_{ob_{skaters}}$, the corresponding referee object $ob_{R_k} \in o^i_{skaters}$.

The ROI for all skaters can be determined by:

$$ROI^i_{skaters} = \Gamma(\bigcup\nolimits_{(\psi^i_l \cap \Lambda^i_{ob_{skaters}} \neq \varnothing)} \psi^i_l) \tag{24}$$

where $\Gamma(\mathcal{L})$ returns an axis-aligned minimum rectangular bounding region for $\mathcal{L}$.

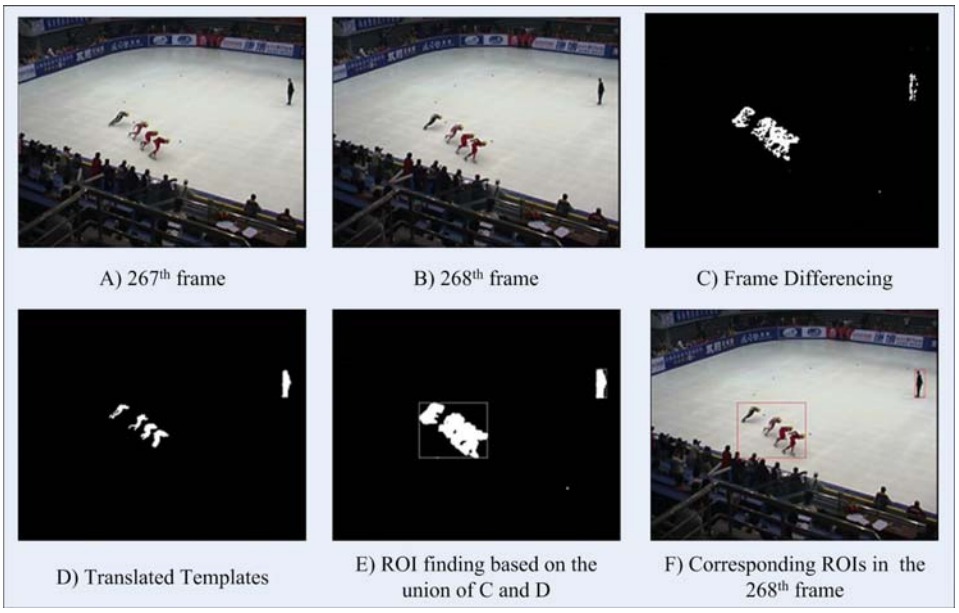

Fig 3.2. An example of ROI generation.

An ROI for the remaining referees can be generated similarly and simply. In the following section, we only focus on $ROI^{i}_{skaters}$. For instance, generated ROIs for 268^{th} frame are illustrated in Figure 3.2F).

3.3 Dynamic Model

In [25], the helmet was a crucial component to determine the position of the skater. Since the real position of the helmet is not on the plane of the rink, using the coordinate system of the rink cannot register the skater correctly. Thus, in [25], a non-linear dynamic model based on adjacent frames and an unscented Kalman filter (UKF) [42] was employed to handle non-linear state estimation. However, such a model cannot predict the real location precisely.

In our system, the helmet is not utilized. As denoted above, $\Gamma(\Omega^{i}_{ob_j})$ represents the axis-aligned minimum rectangular bounding region for ob_j in the *ith* frame. The

coordinate of the center point of the bottom line from the bounding region is considered as the frame position of ob_j, denoted as $p^i_{ob_j}$. By doing so, the approximate feet position is extracted, which most likely lies on the rink. And the real coordinate $r^i_{ob_j}$ for ob_j can be calculated by:

$$(r^i_{ob_j},1)' = \aleph \times (H_i \times (p^i_{ob_j},1)') \tag{25}$$

Based on $r^i_{ob_j}$, a linear dynamic model is proposed here. Denote $X^i_{ob_j} = (D^i_{ob_j}, V^i_{ob_j}, A^i_{ob_j})$ as the state of ob_j in the *ith* frame. $D^i_{ob_j}$ stands for the travelled distance from the first frame to the current *ith* frame, with the unit length equal to 1 meter. $V^i_{ob_j}$ and $A^i_{ob_j}$ represent the velocity and acceleration, respectively. A typical state transition matrix is created as $\begin{bmatrix} 1 & \triangle t & 0 \\ 0 & 1 & \triangle t \\ 0 & 0 & 1 \end{bmatrix}$, where $\triangle t$ equals to $\frac{1}{frame\,rate}$. Since the traveled distance is not easy to observe, the velocity is considered as the measurement and denoted as $\overline{V}^i_{ob_j}$. Assume real coordinates $r^{i-1}_{ob_j}$ and $r^i_{ob_j}$ are known in two consecutive frames. $\overline{V}^i_{ob_j}$ can be calculated by:

$$\overline{V}^i_{ob_j} = \left\| r^{i-1}_{ob_j} - r^i_{ob_j} \right\|^2 \times \triangle r \times \triangle t \tag{26}$$

where $\triangle r$ represents the pixel resolution of the rink coordinate system. As described in the last paragraph of Chapter 2, $\triangle r = 0.05m / pixel$. Both a traditional Kalman filter and UKF can be employed to handle this dynamic model.

How to generate potential real position of object j, based on predicted velocity state $V_{ob_j}^i$, is described as follows.

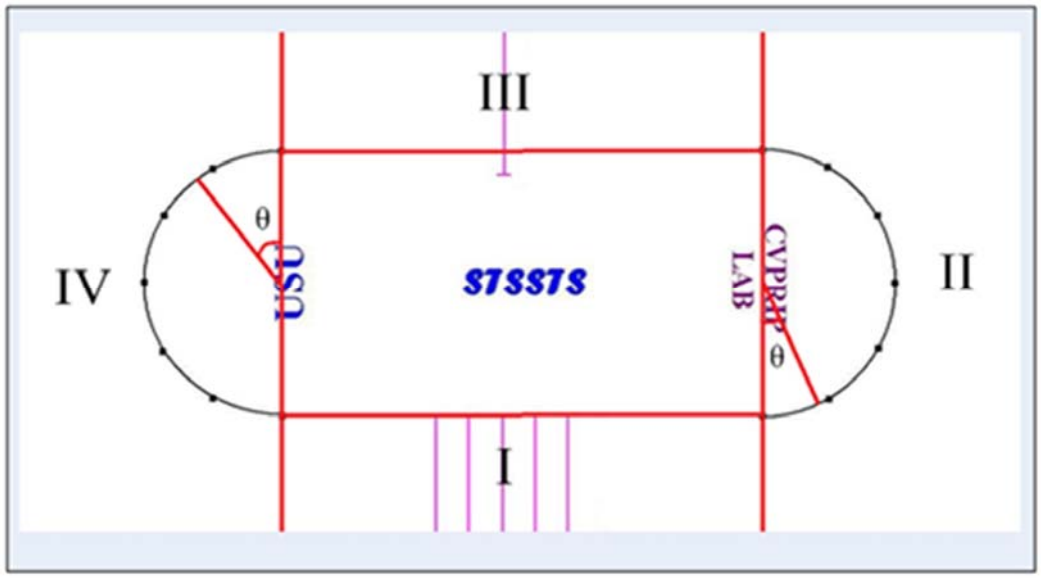

Figure 3.3. Section partitions of the rink coordinate system.

Based on the observation that the skaters always travel counterclockwise on the track, the coordinate system is divided into four regions, as shown in Figure 3.3. In region I, skaters travel from left to right. Conversely, skaters travel from right to left in region III. In regions II and IV, skaters follow the circle curve from bottom to top and from top to bottom, respectively. Denote S as the function to test which region a coordinate belongs to. For instance, $S\left(r_{ob_j}^{i-1}\right) \in$ {I, II, III ,IV} represents the index of the region which ob_j belongs to in frame $i-1$. Based on the current velocity state $V_{ob_j}^i$, the displacement $\triangle d$ can be calculated by $V_{ob_j}^i \times \triangle t / \triangle r$. Assume $r_{ob_j}^{i-1} = (x_{ob_j}^{i-1}, y_{ob_j}^{i-1})$ and the

predicted position $\overline{r}_{ob_j}^{i} = (\overline{x}_{ob_j}^{i}, \overline{y}_{ob_j}^{i})$. If $S(r_{ob_j}^{i-1})$ falls in region I or III, $\overline{r}_{ob_j}^{i}$ can be easily generated by:

$$\begin{cases} \overline{x}_{ob_j}^{i} = x_{ob_j}^{i-1} + \triangle d; \overline{y}_{ob_j}^{i} = y_{ob_j}^{i-1}, & \text{if } S\left(r_{ob_j}^{i-1}\right) = \text{I} \\ \overline{x}_{ob_j}^{i} = x_{ob_j}^{i-1} - \triangle d; \overline{y}_{ob_j}^{i} = y_{ob_j}^{i-1}, & \text{if } S\left(r_{ob_j}^{i-1}\right) = \text{III} \end{cases} \tag{27}$$

If $S(r_{ob_j}^{i-1})$ falls in region II or IV, define $C_l = (x_l, y_l)$ as the center point of the curve with $l = \text{II}$ or IV. Denote $R_{ob_j}^{i-1}$ as the distance of ob_j to $C_{S(r_{ob_j}^{i-1})}$. $R_{ob_j}^{i-1}$ is considered as the radius. As shown in Figure 3.3, radian θ is employed to find $\overline{r}_{ob_j}^{i}$. Define $\theta_{ob_j}^{i-1}$ as the radian for ob_j in frame $i-1$. We obtain:

$$\overline{\theta}_{ob_j}^{i} = \theta_{ob_j}^{i-1} + \triangle d / R_{ob_j}^{i-1} \tag{28}$$

$$\begin{cases} \overline{x}_{ob_j}^{i} = R_{ob_j}^{i-1} \times \sin \overline{\theta}_{ob_j}^{i} + x_{\text{II}}; \overline{y}_{ob_j}^{i} = R_{ob_j}^{i-1} \times \cos \overline{\theta}_{ob_j}^{i} + y_{\text{II}}, & \text{if } S\left(r_{ob_j}^{i-1}\right) = \text{II} \\ \overline{x}_{ob_j}^{i} = R_{ob_j}^{i-1} \times \sin(\overline{\theta}_{ob_j}^{i} + \pi) + x_{\text{IV}}; \overline{y}_{ob_j}^{i} = R_{ob_j}^{i\text{-}1} \times \cos(\overline{\theta}_{ob_j}^{i} + \pi) + y_{\text{IV}}, & \text{if } S\left(r_{ob_j}^{i-1}\right) = \text{IV} \end{cases} \tag{29}$$

The predicted real position $\overline{r}_{ob_j}^{i}$ is obtained through equations 27-29 based on the velocity state $V_{ob_j}^{i}$ and the previous real coordinate $r_{ob_j}^{i-1}$. The corresponding predicted frame position $\overline{p}_{ob_j}^{i}$ can then be calculated:

$$(\overline{p}_{ob_j}^{i}, 1)' = \aleph \times (H_i^{-1} \times (\overline{r}_{ob_j}^{i}, 1)') \tag{30}$$

Referring $\Lambda_{ob_j}^{i}$ as the transformed template of ob_j in the *ith* frame, $q_{ob_j}^{i}$ is defined as the coordinate of the center point of the bottom line from the bounding region $\Gamma(\Lambda_{ob_j}^{i})$.

Denote $\triangle x_{ob_j}^i$ and $\triangle y_{ob_j}^i$ as the horizontal and vertical offsets between $\overline{p}_{ob_j}^i$ and $q_{ob_j}^i$. We obtain $(\triangle x_{ob_j}^i, \triangle y_{ob_j}^i) = \overline{p}_{ob_j}^i - q_{ob_j}^i$. A translation matrix is built:

$$H_{ob_j}^i = \begin{bmatrix} 1 & 0 & \triangle x_{ob_j}^i \\ 0 & 1 & \triangle y_{ob_j}^i \\ 0 & 0 & 1 \end{bmatrix} \tag{31}$$

The translated template of $\Lambda_{ob_j}^i$ is computed by:

$$\overline{\Lambda}_{ob_j}^i = \mathcal{T}(\Lambda_{ob_j}^i, H_{\mathrm{ob_j}}^i) \tag{32}$$

In the following sections, based on $\overline{\Lambda}_{ob_j}^i$, we obtain the evolving template $\Omega_{ob_j}^i$ and the corresponding real position $r_{ob_j}^i$. Then, the measurement can be generated based on $r_{ob_j}^i$ and the previous coordinate $r_{ob_j}^{i-1}$ by equation 26.

3.4 Fuzzy Model

From now on, if there is no specific explanation, all the discussions are based on $ROI_{skaters}^i$ defined in Section 3.2.

The multiple templates evolution algorithm needs a reliable initialization. However, $\overline{\Lambda}_{ob_j}^i$ is not guaranteed to be a suitable choice, which can lead to a failure of the evolution algorithm. Here, we propose a novel fuzzy model to handle such a problem. In classic logic, a pixel can only belong to one object, but in fuzzy logic, a pixel can belong to multiple objects with corresponding memberships. For example, two skaters wearing the same kind of uniform are very close or occluded with each other. Using traditional

logic, one pixel can only belong to one skater and does not affect the clustering result of the other skater. In fuzzy logic, all pixels affect the clustering results of two skaters with different memberships. By doing so, fuzzy logic can achieve a better performance compared with traditional logic. In addition, different from traditional fuzzy c-means algorithm FCM [43, 44], the proposed fuzzy model is based on the evolving template not on traditional cluster centers. The proposed fuzzy model has the advantage to distinguish the shapes of the objects even with similar appearance models.

The rink region is included in the fuzzy model. The set of objects included in $ROI^i_{skaters}$ is expanded as $o^i_{skaters} \bigcup ob_{rink}$.

Define a distance transform function of binary template $\mathcal{G}$ as $\mathfrak{D}(\mathcal{G})$. $\mathfrak{D}$ transforms $\mathcal{G}$ from a binary template to a distance template. For each pixel with 0 value in $\mathcal{G}$, $\mathfrak{D}(\mathcal{G})$ assigns the Euclidian distance between that pixel and the nearest nonzero pixel in $\mathcal{G}$. For each nonzero pixel in $\mathcal{G}$, $\mathfrak{D}(\mathcal{G})$ assigns the distance value as 1. Note that $\overline{\Lambda}^i_{ob_j}$ is a binary template. Define $d^i_{ob_j} = \mathfrak{D}(\overline{\Lambda}^i_{ob_j})$ as the initial distance transform template for ob_j. Then, the membership $u^i_{ob_j}(x, y)$ is defined to represent the degree in which pixel (x, y) belongs to object j. The best result of the proposed fuzzy model can be obtained by minimizing the following objective function:

$$\mathcal{J} = \sum_{(x,y) \in ROI^i_{skaters}} \sum_{ob_j \in o^i_{skaters}} u^i_{ob_j}(x, y)(w_{ob_j}(v^i(x, y)) + d^i_{ob_j}(x, y)) \tag{33}$$

An iterative process is employed to minimize the objective function. $u^i_{ob_j}$ is initialized by $\overline{\Lambda}^i_{ob_j}$. And $u^i_{ob_j}$ can be iteratively updated by the following equations:

$$\begin{cases} \overline{\Lambda}^i_{ob_j}(x,y)=1, & \text{if } ob_j = \underset{ob_l}{argmax}\, u^i_{ob_l}(x,y) \\ \overline{\Lambda}^i_{ob_j}(x,y)=0, & \text{otherwise} \end{cases} \tag{34}$$

$$\overline{d}^i_{ob_j}(x,y) = \mathfrak{D}(\overline{\Lambda}^i_{ob_j}) \tag{35}$$

$$\overline{u}^i_{ob_j}(x,y) = \frac{1}{\sum_{ob_l \in o^i_{skaters}} \left(\frac{w_{ob_j}(v^i(x,y)) + \overline{d}^i_{ob_j}(x,y)}{w_{ob_l}(v^i(x,y)) + \overline{d}^i_{ob_j}(x,y)} \right)} \tag{36}$$

Note that since the denominator of the updated membership $\overline{u}^i_{ob_j}(x,y)$ is always larger than or equal to 1, the membership $\overline{u}^i_{ob_j}(x,y) \in [0,1]$. Calculate measurement η by:

$$\eta = \max_{(x,y) \in ROI^i_{skaters}, ob_j \in o^i_{skaters}} | u^i_{ob_l}(x,y) - \overline{u}^i_{ob_j}(x,y) | \tag{37}$$

The iteration stops when $\eta < \varepsilon$, where ε is a termination criterion between 0 and 1. Set $\varepsilon = 0.1$ was determined by experiments in this dissertation. Another criterion is the maximum number of iterations, which is 5. If the iteration is not terminated, swamp membership by equation 38 and start a new iteration.

$$u^i_{ob_l}(x,y) = \overline{u}^i_{ob_j}(x,y) \tag{38}$$

Now, an initial template $\xi^i_{ob_j}$ for template evolution can be generated according to $u^i_{ob_l}$ and the transformed template $\Lambda^i_{ob_j}$ which is not changed by the fuzzy model. Denote $T^i_{ob_j} = \Gamma(\Lambda^i_{ob_j})$ as a template filter. Thus, $T^i_{ob_j}$ represents the axis-aligned minimum rectangular bounding region of ob_j with the transformed shape from frame $i-1$. Note

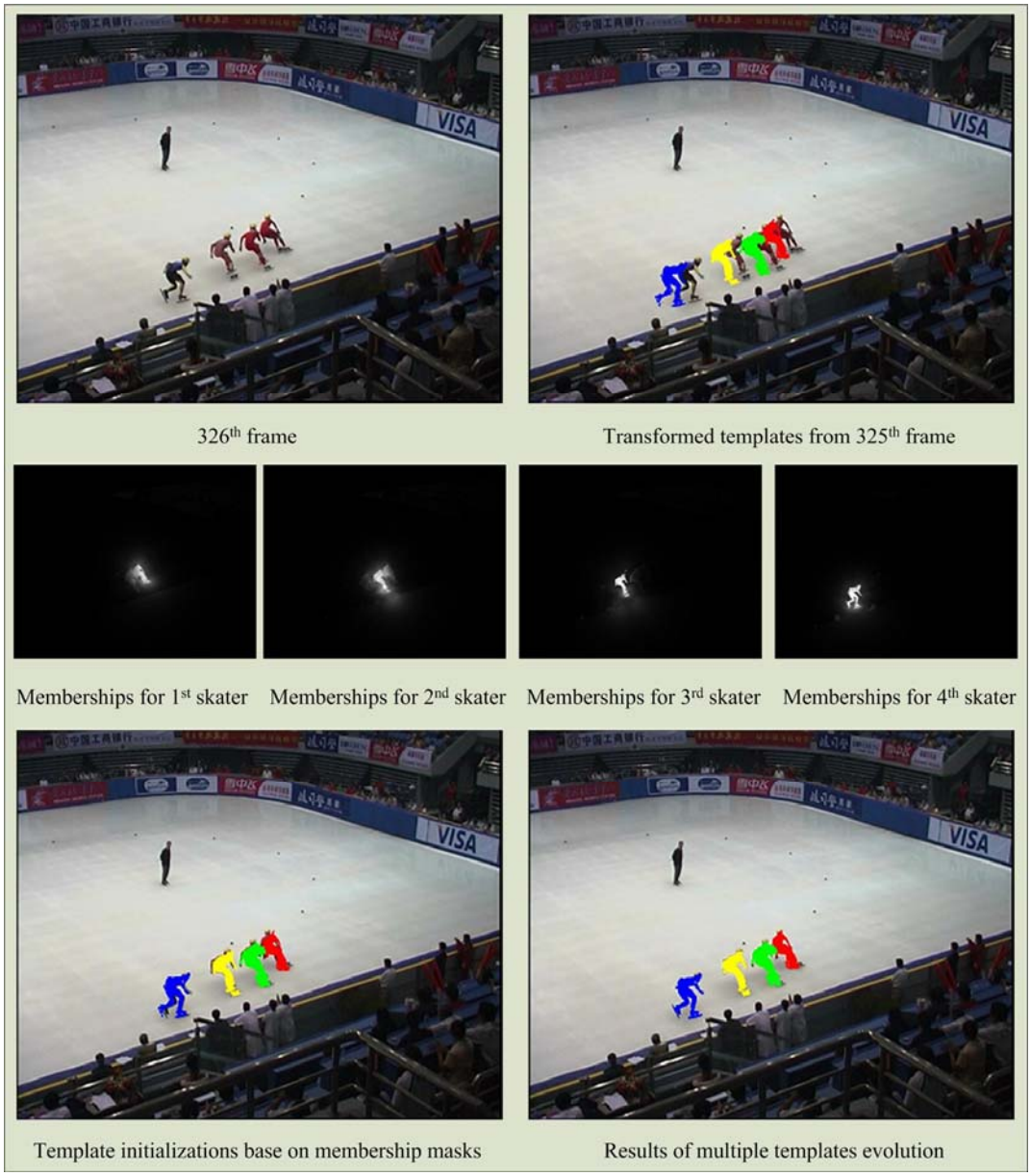

Figure 3.4. Illustration of the capability of the proposed fuzzy model.

that after the termination of the fuzzy model, $\overline{\Lambda}^i_{ob_j}$ encodes the most likely position for $T^i_{ob_j}$. Let $\Phi^i_{ob_j} = \Gamma(\overline{\Lambda}^i_{ob_j})$ represent the candidate region of the center position of $T^i_{ob_j}$.

Perform a convolution operation (operator denoted as *) on membership mask $u^i_{ob_l}$ with the center of the convolved mask $T^i_{ob_j}$ locating only in the region of $\Phi^i_{ob_j}$:

$$G^i_{ob_j} = T^i_{ob_j} * u^i_{ob_l} \tag{39}$$

The position $(\tilde{x}, \tilde{y}) = \arg\max_{(x,y)\in\Phi^i_{ob_j}} G^i_{ob_j}(x, y)$ is considered as the optimum position for $T^i_{ob_j}$. The initial template $\xi^i_{ob_j}$ is obtained by translating the center of $T^i_{ob_j}$ to position $(\tilde{x}, \tilde{y})$.

As illustrated in Figure 3.4, even the transformed templates are quite off target. In particular, the first two players wear the same kind of uniform, and the transformed template for the first player is in the middle of the two players. However, fuzzy memberships can still locate the correct regions for all skaters successfully. The initial templates $\xi^i_{ob_j}$ for ob_j is generated superbly based on memberships.

3.5 Multiple Template Evolution

The initial template $\xi^i_{ob_j}$ is generated as described in the above section. Now, a multiple-template evolution approach is proposed to obtain the optimum evolving template $\Omega^i_{ob_j}$ for ob_j in the *ith* frame. To simplify the notation, the set of objects included in $ROI^i_{skaters}$ is redefined by o^i and $ob_{rink} \notin o^i$.

Notice that the object templates might overlap with each other when they are initialized or during evolution steps. Define a monitor structure $\mathcal{M}^i$, which is initialized by:

$$\mathcal{M}^i(x,y)=\left\{ob_j|\xi^i_{ob_j}(x,y)=1\right\} \tag{40}$$

A level set-based active contour algorithm [45] is employed and improved to realize multiple templates evolution. A weighted GMM template w_{ob_j} is considered as the input image for ob_j. Compared with RGB space, w_{ob_j} is more homogeneous since the uniforms of the skaters may include several kinds of colors.

Define level set function $\phi^i_{ob_j}$ for object j as:

$$\begin{cases}(x,y)\in \text{ contour of } ob_j & \text{if } \phi^i_{ob_j}(x,y)=0\\ (x,y) \text{ is inside } ob_j & \text{if } \phi^i_{ob_j}(x,y)>0\\ (x,y) \text{ is outside } ob_j & \text{if } \phi^i_{ob_j}(x,y)<0\end{cases} \tag{41}$$

$\phi^i_{ob_j}$ can be initialized based on $\mathcal{M}^i$ using the similar strategy of obtaining distance transform mask $d^i_{ob_j}$. For details, please refer to [45].

Define $c_1(\phi^i_{ob_j})$ and $c_2(\phi^i_{ob_j})$ as the inside average and outside average of the contour for ob_j. Since there are many other objects outside the contour of ob_j, using the method in [45] does not achieve the optimum constant value to represent the characteristic of the background as indicated in [46]. However, by using monitor $\mathcal{M}^i$ to exclude possible object points, this problem can be avoided; c_1 and c_2 are calculated as:

$$\begin{cases}c_1\left(\phi^i_{ob_j}\right)=\text{mean}\left\{w_{ob_j}(x,y)|\phi^i_{ob_j}(x,y)\geq 0\right\}\\ c_2\left(\phi^i_{ob_j}\right)=\text{mean}\left\{w_{ob_j}(x,y)|\phi^i_{ob_j}(x,y)<0\wedge\mathcal{M}^i(x,y)=\varnothing\right\}\end{cases} \tag{42}$$

The energy function for multiple templates evolution algorithm can be defined as:

$$F^i = \sum_{ob_j \in o^i} (\mu L_{ob_j} + \lambda_1 \sum_{ob_j \in \mathcal{M}^i(x,y) \wedge \phi^i_{ob_j}(x,y) \geq 0} \left| w_{ob_j}(x,y) - c_1\left(\phi^i_{ob_j}\right) \right| + \lambda_2 \sum_{ob_j \in \mathcal{M}^i(x,y) \wedge \phi^i_{ob_j}(x,y) < 0} \left| w_{ob_j}(x,y) - c_2\left(\phi^i_{ob_j}\right) \right|) \quad (43)$$

where L_{ob_j} represents the length of the contour of ob_j. Parameters λ_1 and λ_2 are set to 1, and the smoothing constrain μ is set to 0.2 determined by experiments.

Add one more constraint on energy function such that no contour is occluded by others. Without this constraint, the evolution is likely to over evolve the contours of the skaters wearing similar uniforms, especially, when they are close to each other. However, there are often overlapped skaters on the rink, and this is certainly one of the challenging issues to be faced. To overcome this problem, besides the global monitor structure $\mathcal{M}^i$, an occluded template $O^i_{ob_j}$ is defined for object j. O^i limits the cardinality (denoted by $|\cdot|$) of the element in $\mathcal{M}^i$ to be less than or equal to 1, which means $|\mathcal{M}^i(x,y)| \leq 1$. Meanwhile, O^i keeps the occlusion information.

Since the initial templates have a great chance to match the targets, the occluded objects in $\mathcal{M}^i$ are most likely caused by real occlusions among objects. Based on another observation that the skater closest to the camera will occlude the others, we employ the frame position $p^{i-1}_{ob_j}$, which is defined in Section 3.3, to determine the relationships among the occluded objects. Let $p^{i-1}_{ob_j} = (x^{i-1}_{ob_j}, y^{i-1}_{ob_j})$. Now, $O^i_{ob_j}$ can be obtained by equation 44 based on $\mathcal{M}^i$:

$$\begin{cases} O_{ob_j}^{i}(x,y)=1, & \text{if } y_{ob_j}^{i-1} \neq \max\limits_{ob_l \in \mathcal{M}^i(x,y)} \{y_{ob_j}^{i-1}\} \wedge ob_j \in \mathcal{M}^i(x,y) \\ O_{ob_j}^{i}(x,y)=0, & \text{if } y_{ob_j}^{i-1} = \max\limits_{ob_l \in \mathcal{M}^i(x,y)} \{y_{ob_j}^{i-1}\} \vee ob_j \notin \mathcal{M}^i(x,y) \end{cases} \tag{44}$$

And $\mathcal{M}^i$ can be updated based on $O_{ob_j}^{i}$ so that the cardinality of the element in $\mathcal{M}^i$ is less than or equal to 1 by equation 45.

$$\mathcal{M}^i(x,y)=ob_j, \text{ if } ob_j \in \mathcal{M}^i(x,y) \wedge O_{ob_j}^{i}(x,y)=0 \tag{45}$$

Define $\mathcal{M}_k^i$ as the updated M^i and $O^{i,k}$ as the updated O^i in the kth iteration, respectively. The iterative templates evolution process to minimize the energy function with occlusion constrains is described as follows:

Initial step: Set $k=0$ as the initial iteration number and $k_{max}=20$ as the max iteration number. Initialize $\mathcal{M}_0^i$ by equation 40. Initialize $O^{i,0}$ by equation 44, and update $\mathcal{M}_0^i(x,y)$ by equation 45. Initialize level sets function $\phi_{ob_j}^{i,0}$ for all ob_j based on $\mathcal{M}_0^i(x,y)$. Set k =1, start iteration.

Step 1: Compute $c_1(\phi_{ob_j}^{i,k-1})$ and $c_2(\phi_{ob_j}^{i,k-1})$ for all ob_j by equation 42.

Step 2: Update each $\phi_{ob_j}^{i,k}$ based on [45], using $c_1(\phi_{ob_j}^{i,k-1})$, $c_2(\phi_{ob_j}^{i,k-1})$ and w_{ob_j}.

Step 3: Update $\mathcal{M}_k^i$ by:

$$\begin{cases} \mathcal{M}_k^i(x,y)=\mathcal{M}_k^i(x,y) \bigcup ob_j, & \text{if } \phi_{ob_j}^{i}(x,y) \geq 0 \wedge |\mathcal{M}_{k-1}^i(x,y)|=0 \\ \mathcal{M}_k^i(x,y)=\mathcal{M}_k^i(x,y)-ob_j, & \text{if } \phi_{ob_j}^{i}(x,y) < 0 \wedge \mathcal{M}_{k-1}^i(x,y)=ob_j \end{cases} \tag{46}$$

Update $O^{i,k}$ by equation 44 and update $\mathcal{M}_n^i$ by equation 45.

Step 4: Check whether $\mathcal{M}_k^i(x,y)$ is equal to $\mathcal{M}_{k-1}^i(x,y)$. If not, check whether $k \le k_{max}$. If not, set $k = k+1$. Go to step 1. Otherwise, stop.

The optimum evolving template $\Omega_{ob_j}^i$ can be generated by:

$$\begin{cases} \Omega_{ob_j}^i(x,y) = 1, & \text{if } ob_j \in \mathcal{M}_k^i(x,y) \\ \Omega_{ob_j}^i(x,y) = 0, & \text{if } ob_j \notin \mathcal{M}_k^i(x,y) \end{cases} \tag{47}$$

Based on $\Omega_{ob_j}^i$, the real coordinate $r_{ob_j}^i$ can be obtained by equation 25.

Note that $\Omega_{ob_j}^i$ does not include the occluded template $O_{ob_j}^i$. Because the occluded part is indeed invisible in the frame, it affects the predicted position for the initial template described in Section 3.4. However, we still need $O_{ob_j}^i$ to handle occlusion issues. Here, we expand $O_{ob_j}^i$ at the end of evolution process by:

$$O_{ob_j}^i = \Omega_{ob_j}^i(x,y) \bigcup O_{ob_j}^{i,k} \tag{48}$$

Assume we have the occluded template $O_{ob_j}^{i-1}$ in frame $i-1$. We can obtain a transformed occluded template in frame i by $\bar{O}_{ob_j}^i = \mathcal{T}(O_{ob_j}^{i-1}, H_{i-1,i})$. Denote $\bar{\xi}_{ob_j}^i = \Gamma(\bar{O}_{ob_j}^i)$ as an axis-aligned minimum rectangular bounding region of $\bar{O}_{ob_j}^i$. The center of $\bar{\xi}_{ob_j}^i$ is set as the center of $\xi_{ob_j}^i$, and $\xi_{ob_j}^i$ in equation 40 is replaced by $\bar{\xi}_{ob_j}^i$. Since $\xi_{ob_j}^i \subset \bar{\xi}_{ob_j}^i$, the optimum position for $\xi_{ob_j}^i$ should also be the optimum position for $\bar{\xi}_{ob_j}^i$. By doing so, we use $\Omega_{ob_j}^{i-1}$ to find the optimum position for ob_j in frame i before evolution process while use $O_{ob_j}^{i-1}$ to handle occlusion problems during evolution process.

Figure 3.4 shows example results of multiple-template evolution algorithm. Note that the initial templates obtained by the proposed fuzzy model are slightly different from the real shapes of the target objects. However, the evolving templates generated by multiple templates evolution algorithm match the target objects perfectly.

More complex cases are illustrated in Figure 3.5. There are several kinds of occlusions such as group situations, occlusions between skaters wearing different kind of uniforms, and occlusions between skaters wearing same kind of uniforms. The results demonstrate that the proposed algorithm can handle all of the above cases successfully. Finally, Figure 3.6 shows the flowchart and explanation for the entire tracking system.

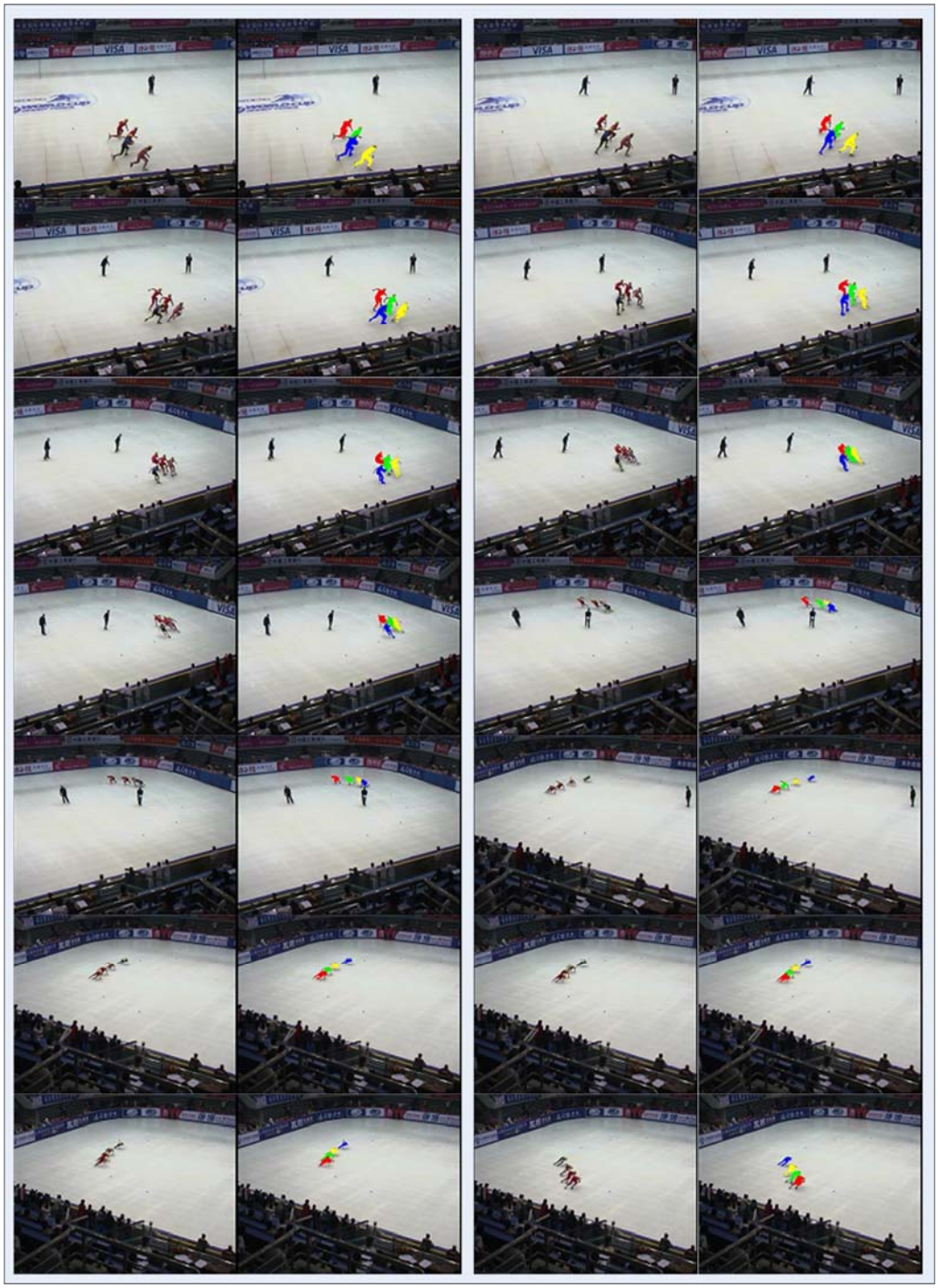

Fig 3.5. Results of multiple templates evolution algorithm. The first column displays original frames indexed as 44, 68, 88, 105, 133, 246, and 250; and the second column displays the corresponding resulting frames. The third column displays original frames indexed as 58, 81, 97, 128, 241, 248 and 261; and the fourth column displays the corresponding resulting frames.

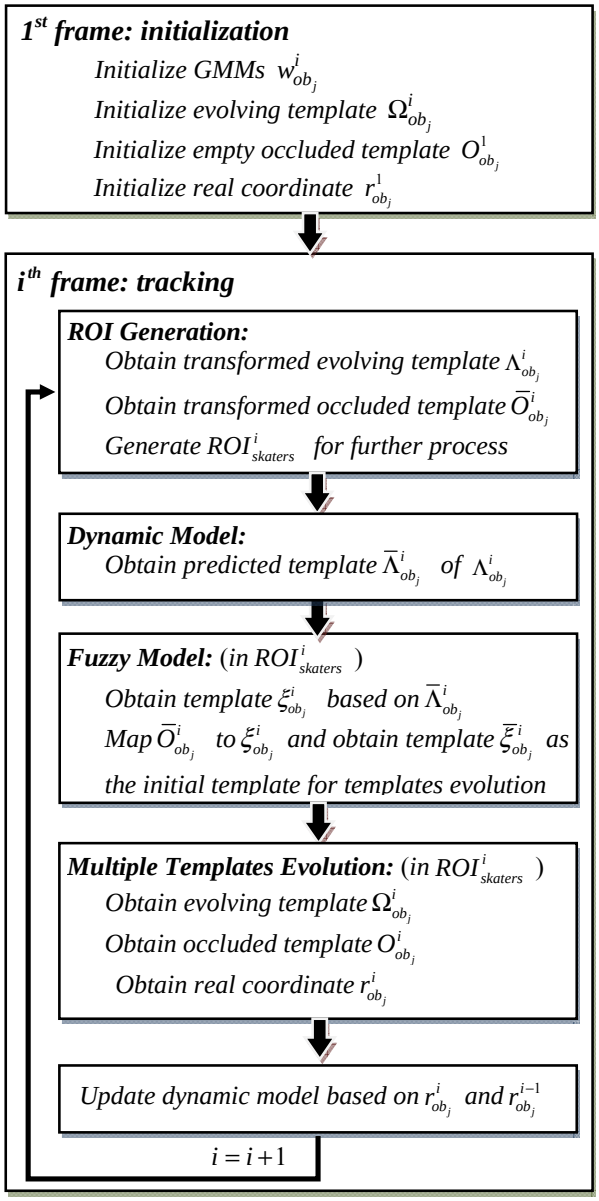

Figure 3.6. Flowchart and explanation for tracking system.

CHAPTER 4

EXPERIMENTS AND RESULTS

Two testing videos are obtained from 2005/2006 BOSIDEN ISU International Skating Union (ISU) world cup short track speed skating. One is the men's 500m finals, which has 1071 frames with frame rate $25\ frames/\sec$. There are four skaters in the video, two skaters of which are from China. The other two are from Korea and Canada, respectively. Denote the first video as video 1. The other one is the women's 500m semi-finals which has 1156 frames with the same frame rate as video 1. Four skaters are in the video, two from Canada, and other two are from China and Russia, respectively. Denote the second video as video 2. Skaters coming from the same country wear the same kind of uniforms. Table 4.1 lists the number of mixture Gaussians assigned to the skaters of different countries.

Table 4.1. Number of Mixture Gaussians for Skaters from Different Countries.

	China	Canada	Korean	Russia
#.mixtures	2	2	3	3

The real coordinate $r_{ob_j}^i$ is the basis for the 2D reconstruction. The skaters are modeled as 2D balls with different colors. The size of the 2D reconstructed virtual rink is 1200×600. By mapping the real rink with $60m\times30m$ to the 2D reconstructed virtual rink, we obtain the resolution of the virtual rink is $5cm/pixel$.

4.1 Tracking and Reconstruction Results

As in many tracking systems, to clearly show the tracking result and to keep the target visible, an axis-aligned minimum rectangular bounding box is employed to highlight the detected object. The bounding box of object j in the ith frame can be generated by the evolving template $\Omega_{ob_j}^{i}$.

In Figure 4.1, twelve sample frames that represent the effectiveness of the proposed algorithm are selected from the 603rd frame to the 667th frame in video 1. During this period, the leading position is switched between two Chinese skaters. Skater A trailed Skater B at the beginning. Skater B continuously accelerated, and finally surpassed Skater A at the 653rd frame. Note that serious occlusion happens between the two Chinese skaters from frames 634th to 659th. Even though they wear identical uniforms, they still can be tracked successfully. Another observation is that the last skater is dramatically overlapped with the advertising board from the 621st frame to the 653rd frame. Since the audience section is eliminated, much of the last skater cannot be detected. However, some parts of his legs are still tracked correctly. After the 653rd frame, when the last skater moves out from the advertising board region, he is recovered successfully. This observation shows the ability of the system to handle the problem of dramatic shape change. Furthermore, all referees are also tracked successfully. Since referees move slowly in relation to the athletes, the proposed system tracks referees easily without using the dynamic model.

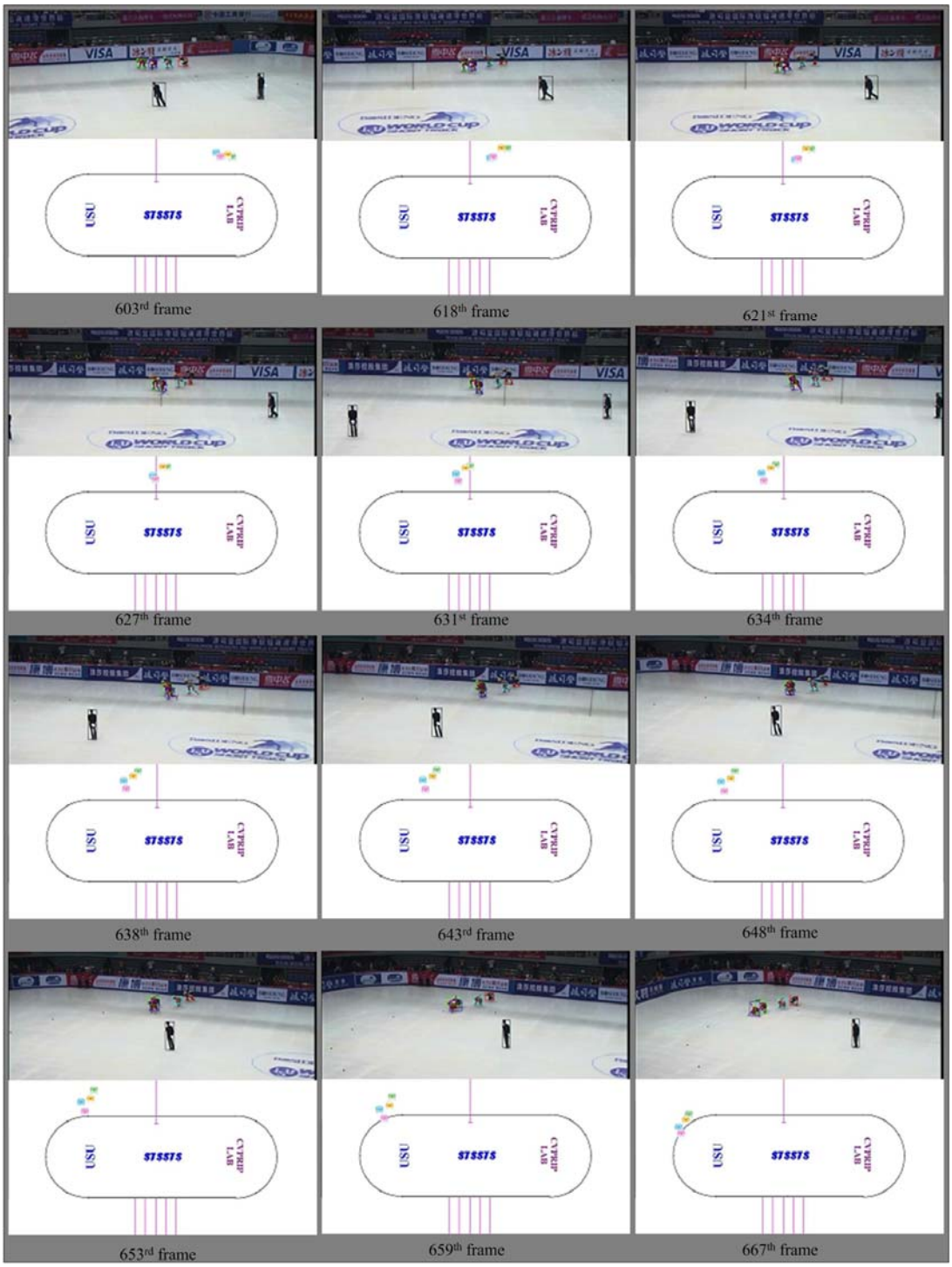

Figure 4.1. Tracking and reconstruction results for 12 frames in video 1. For each frame, the top half is the tracking result, and the bottom part is the corresponding reconstruction result.

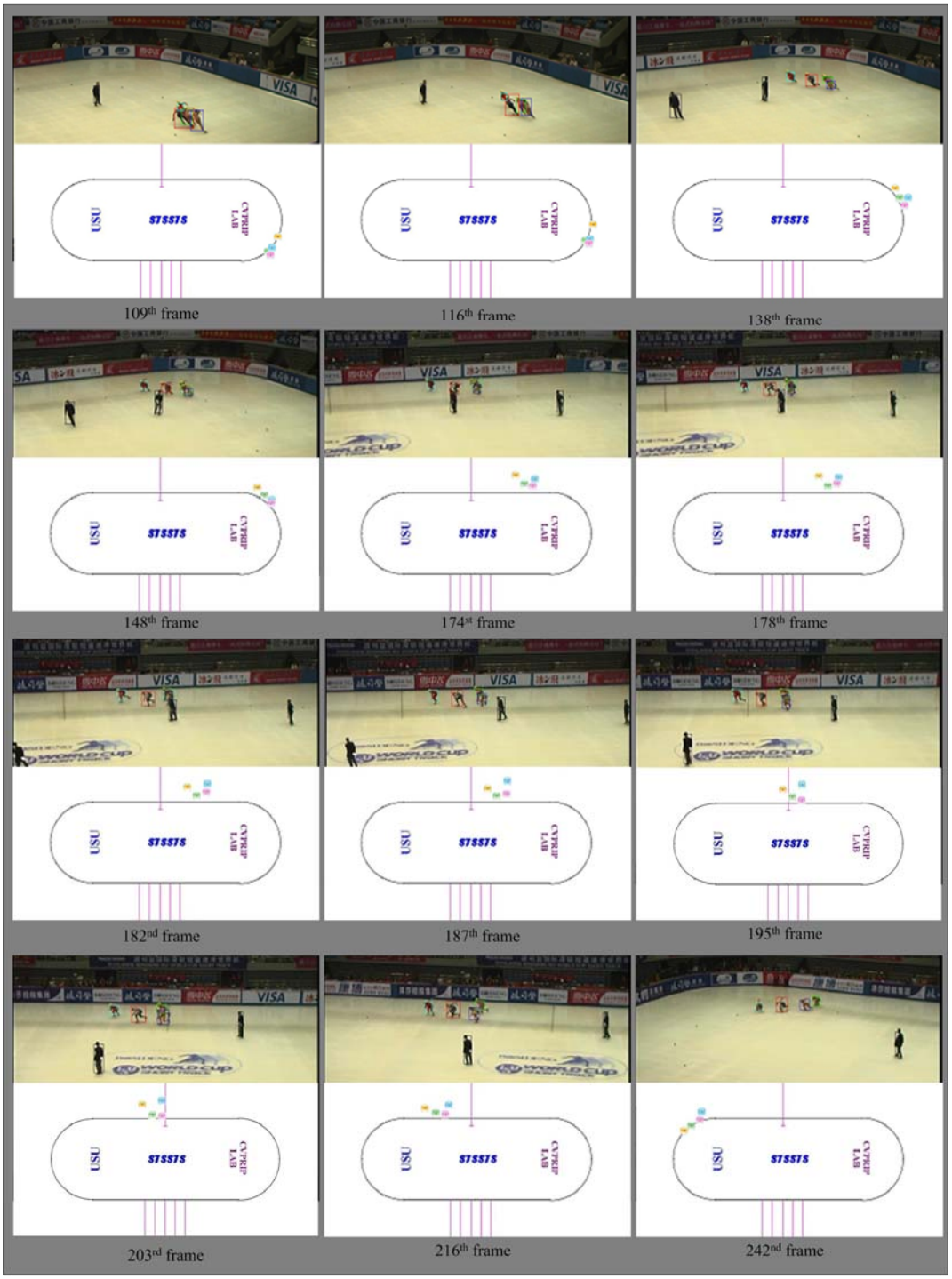

Figure 4.2. Tracking and reconstruction results for 12 frames in video 2. For each frame, the top half is the tracking result, and the bottom part is the corresponding reconstruction result

In Figure 4.2, another twelve frames that also represent the effectiveness of the proposed algorithm are selected from the 109th frame to the 242nd frame in video 2. Starting from the 109th frame, serious group occlusion happens. From the 109th frame to the 203rd frame, there are serious occlusions between the two Canadian skaters. At the 216th frame, they are separated, and the lead position is switched between the two Canadian skaters. Also, in 174th frame and 182nd frame, there are occlusions between skaters and a referee. As shown in Figure 4.2, the proposed system handles all these situations very well.

On the other hand, the 2D reconstruction animations displayed by the bottom half of each frame in Figures 4.1 and 4.2 precisely mark the real positions for the skaters frame by frame. These data are invaluable for sports experts and coaches.

4.2 Trajectory and Velocity Analysis

In Figure 4.3, the trajectories of two female skaters are extracted. Different colors represent different laps, namely, the first lap is red; the second green; the third blue; the fourth yellow; and last half-lap black.

The top trajectory belongs to the skater in first place, while the bottom one belongs to the skater in last place. The trajectories of the two skaters are almost the same. However, the angle-of-entering-the-curve of the first skater is slightly smaller than that of the last skater. This indicates that the curve technique of the first skater is better than that of the last skater. Curve technique is crucial for short track skaters. In fact, generally speaking, a skater's curve technique determines whether the skater is an elite athlete.

Thus, the ability to visually track curve technique through trajectory analysis is highly desirable.

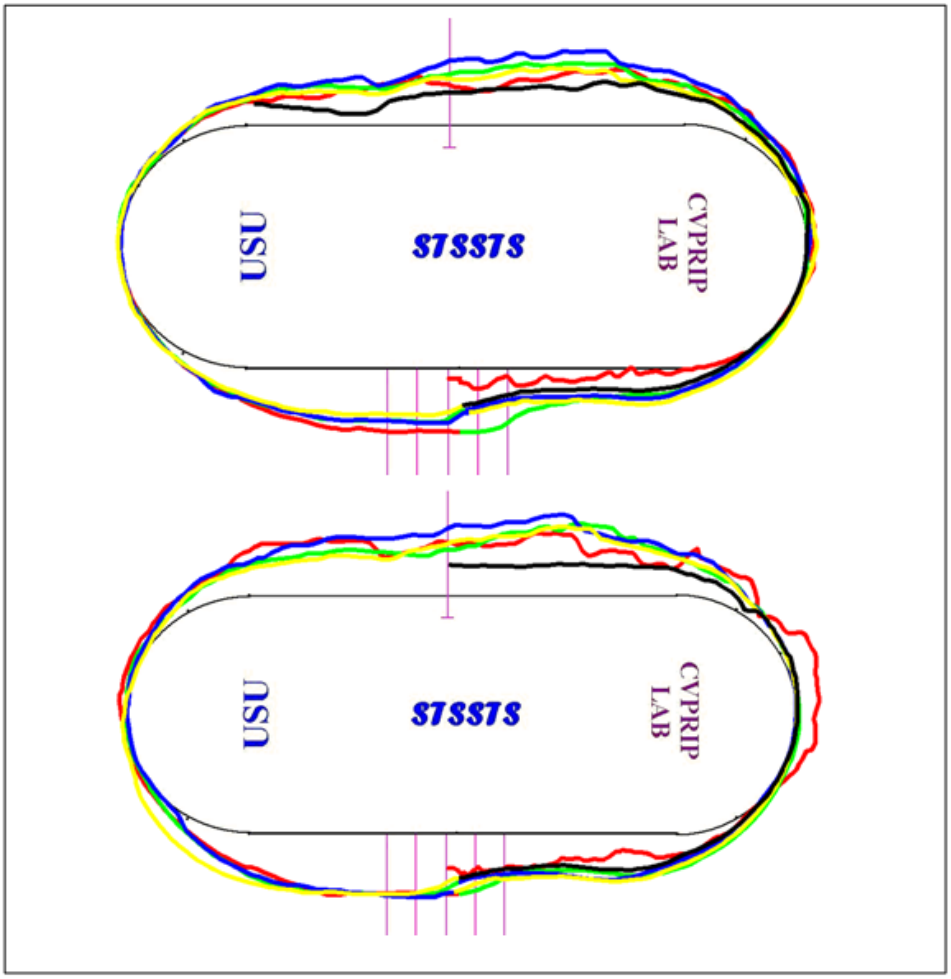

Figure 4.3. Trajectories of two female skaters in video 2.

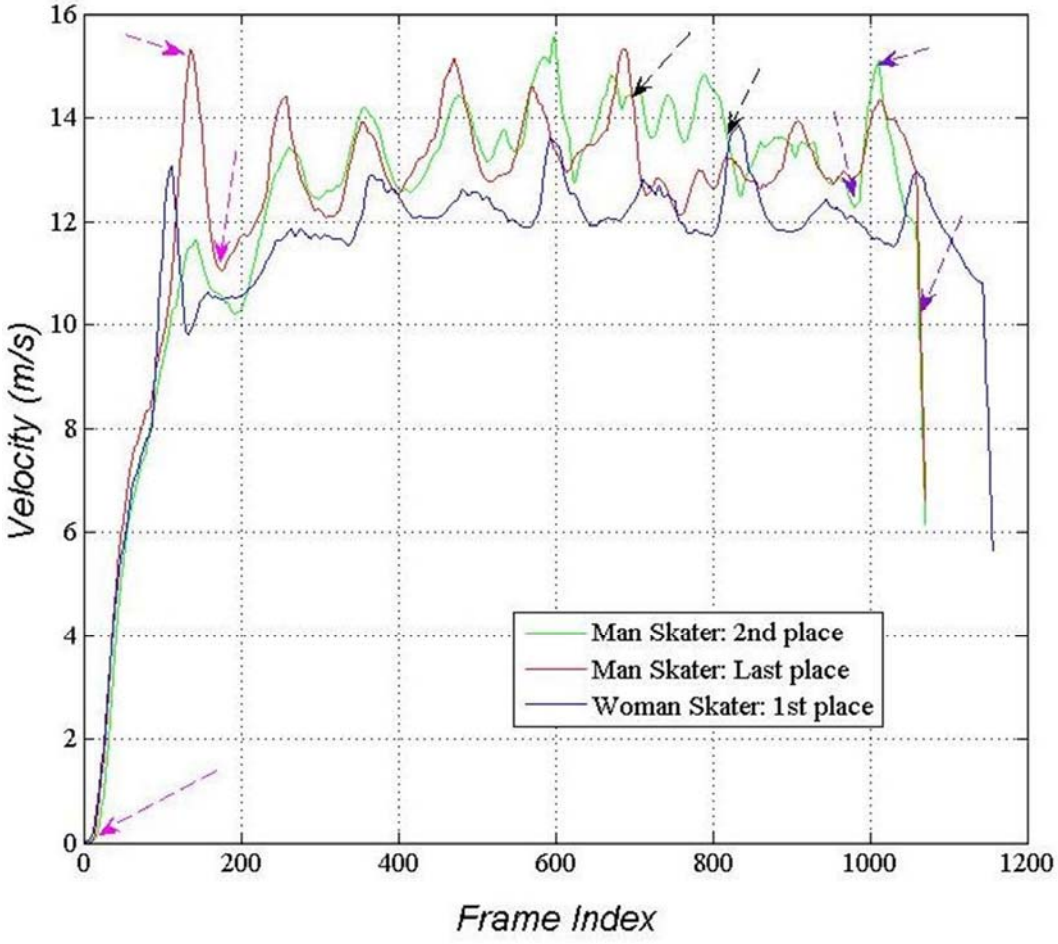

Figure 4.4. Velocity curves for two male skaters and one female skaters

In Figure 4.4, three velocity curves are displayed: two male skaters, one in second place and the other in last place; and the female skater in first place. Several interesting observations can be made according to these velocity curves.

1) In a 500m competition, there are in total nine rink curves. Crowding often happens at the rink curves, which in all likelihood results in a decrease in velocity. Note that there is a similar number of the peak for each curve.

2) The period, pointed out by the three pink arrows, is the start of the race for the male skater in the last place. During this period, the velocity grows dramatically. When he achieves the first rink curve, however, his velocity decreases. A similar phenomenon can be observed of the other skaters' curves.

3) After the first rink curve, the velocity keeps at a certain level with slight changes.

4) Even the fastest female skater is slower than the slowest male skater according to these two videos. Even so, the female skater maintains first place in her race. Crowding rarely affects her. The velocity curve of the female skater is much smoother than that of the two male skaters.

5) The three blue arrows indicate the sprint stages for the male skater in second place. After reaching the finish line, the velocity decreases dramatically. A similar phenomenon can be observed from the curves of the other skaters.

6) The male skater in second place surpasses the male skater in fourth place during the period pointed out by the two black arrows. Before this period, the male skater in the second place is consistently behind the male skater in fourth place. But after this period, he is never surpassed again by his opponent.

4.3 Quantitative Analysis

For all skaters in each frame, the axis-aligned minimum rectangular bounding boxes are outlined manually. The coordinate of the center point of the bottom line of the bounding box of object j in ith frame is automatically extracted and denoted as $\hat{p}^{i}_{ob_j}$.

The real coordinate $\hat{r}_{ob_j}^i$ is obtained by equation 25 and is considered as the ground truth for the quantitative analysis.

Next, the mean (avg) and standard deviation (std) are employed as two metrics of the system and can be calculated as:

$$avg_{ob_j} = \frac{\sum_{i=1}^{n} \mathcal{D}_{ob_j}^i}{n} \tag{48}$$

$$std_{ob_j} = \sqrt{\frac{\sum_{i=1}^{n} (\mathcal{D}_{ob_j}^i - avg_{ob_j})^2}{n}} \tag{49}$$

where $\mathcal{D}_{ob_j}^i$ is the Euclidian distance between the ground truth $\hat{r}_{ob_j}^i$ and the tracking result $r_{ob_j}^i$. The smaller the avg_{ob_j}, the more accurate the system is. The smaller the std_{ob_j}, the more robust the system is. The results for all skaters in the two videos are shown in Table 4.2. Note that both avg_{ob_j} and std_{ob_j} are less than 2.5 pixels in the rink coordinate system and less than 13 cm in the real world. Thus, the proposed system is quite accurate and robust.

Table 4.2. Tracking Errors for All Skaters in Pixels and Meters (Inside the Brackets).

	Male1	Male2	Male3	Male4
avg_{ob_j}	1.31(0.07)	0.89(0.04)	1.45(0.07)	1.23(0.06)
std_{ob_j}	2.31(0.12)	1.25(0.06)	2.32(0.12)	1.40(0.07)
	Women1	Women2	Women3	Women4
avg_{ob_j}	1.58(0.08)	0.89(0.04)	0.78(0.04)	1.60(0.08)
std_{ob_j}	1.64(0.08)	1.03(0.05)	1.03(0.05)	1.74(0.09)

CHAPTER 5

CONCLUSION

In this dissertation, we propose a novel computer vision system STS2 to track multiple high-speed deformable skaters in short track speed skating video captured by a single panning camera. Short track speed skating video suffers several open-ended, challenging problems. To overcome such problems, we propose a novel and effective approach consisting of several new algorithms/models:

1) An efficient automatic key frame finding algorithm is proposed to obtain the homography transforming each frame to a real world rink coordinate system. By doing so, the motion of the camera is removed.
2) A dynamic model is introduced based on rink parameters to make the motion of the skaters linearly predictable.
3) An evolving template is employed to model non-rigid skaters.
4) A novel fuzzy model is proposed to predict the optimum initial positions for templates evolution algorithm.
5) A novel multiple-template evolution algorithm is proposed. Combined with a fuzzy model, occlusions among objects are successful handled. This holds true even for objects having a similar appearance.
6) An automatic ROI-generation algorithm is implemented to locate the target rapidly.

The experimental results demonstrate that the proposed system solves the challenging problems very effectively.

The end results of the system, including trajectories, velocity analysis and 2D reconstruction animations, are invaluable to broadcasters and coaches alike. The proposed system has the potential to be applied to many other related areas.

The videos of complete tracking and 2D reconstruction animation can be downloaded from ***http://cvprip.cs.usu.edu/wang/STSSTS***.

REFERENCES

1. Misu, T., Naemura, M., Zheng, W., Izumi, Y., and Fukui, K. Robust tracking of soccer players based on data fusion. In *Sixteenth International Conference on Pattern Recognition*, 2002, 556–561.
2. Vandenbroucke, N., Macaire, L., and Postaire, J.G. Color image segmentation by pixel classification in an adapted hybrid color space. Applications to soccer image analysis. *Computer Vision and Image Understanding 90*, 2 (2003), 190–216.
3. Needham, C.J. and Boyel, R.D. Tracking multiple sports players through occlusion, congestion and scale. In *Twelfth British Machine Vision Conference*, 2001, 93–102.
4. Liu, J., Tong, X., Li, W., Wang, T., Zhang, Y., and Wang, H. Automatic player detection, labeling and tracking in broadcast soccer video. *Pattern Recognition Letters 30* 2 (January 2009), 103-113.
5. D'Orazio, T. and Leo, M. A review of vision-based systems for soccer video analysis. *Pattern Recognition 43*, 8 (2010), 911-2926.
6. Pingali, G.S., Jean, Y., and Carlborn, I. Real time tracking for enhanced tennis broadcasts. In *International Conference on Computer Vision and Pattern Recognition*, 1998, 260–265.
7. Yan, F., Kostin, A., Christmas, W.J., and Kittler, J. A novel data association algorithm for object tracking in clutter with application to tennis video analysis. In *International Conference on Computer Vision and Pattern Recognition*, 2006, 634–641.

8. Pers, J., Vuckovic, G., Kovacic, S., and Dezman, B. A low-cost real-time tracker of live sport events. In *Second International Symposium on Image and Signal Processing and Analysis*, 2001, 362–365.

9. Pers, J., Bon, M., Kovacic, S., Sibila, M., and Dezman, B. Observation and analysis of large-scale human motion. *Human Movement Science 21*, 2 (2002), 295–311.

10. Jug, M., Pers, J., Dezman, B., and Kovacic, S. Trajectory based assessment of coordinated human activity. In *Proceedings of the Third International Conference on Computer Vision Systems*, 2003, 534–543.

11. Intille, S.S. and Bobick, A.F. Tracking using a local closed-world assumption: tracking in the football domain. Technical Report 296, MIT Media Lab Perceptual Computing Group, 1994.

12. Okuma, K. *Automatic Acquisition of Motion Trajectories: Tracking Hockey Players*. Master's Thesis, The University of British Columbia, 2003.

13. Xing, J., Ai, H., Liu, l., and and Lao, S. Multiple player tracking in sports video: A dual-mode two-way Bayesian inference approach with progressive observation modeling, *IEEE Trans. Image Processing 20*, 6 (2011), 1652–1667.

14. Haritaoglu, I., Harwood, D., and Davis, L.S. W4: Real-time surveillance of people and their activities. *IEEE Trans. on Pattern Analysis and Machine Intelligence 22*, 8 (2000), 809–830.

15. Wren, C.R., Azarbayejani, A., Darrell, T., and Pentland, A. Pfinder: Real-time tracking of the human body. *IEEE Trans. on Pattern Analysis and Machine Intelligence 19*, 7 (1997), 780–785.

16. Intille, S., Davis, J., and Bobick, A. Real-time closed-world tracking. In *International Conference on Computer Vision and Pattern Recognition*, 1997, 697–703.

17. Masoud, O. and Papanikolopoulos, N.P. A novel method for tracking and counting pedestrians in real-time using a single camera. *IEEE Trans. on Pattern Analysis and Machine Intelligence 50*, 5 (2001), 1267–1278.

18. Rigoll, G., Breit, H., and Wallhoff, F. Robust tracking of persons in real-world scenarios using a statistical computer vision approach. *Image and Vision Computing 22*, 7 (2004) 571–582.

19. Bradski, G.R. Computer vision face tracking for use in a perceptual user interface. In *IEEE Workshop on Applications of Computer Vision*, 1998, 214–219.

20. Chen, Y., Rui, Y., and Huang, T.S. Multicue HMM-UKF for real-time contour tracking. *IEEE Trans. on Pattern Analysis and Machine Intelligence 28*, 9 (2006), 1525–1529.

21. Comaniciu, D., Ramesh, V., and Meer, P. Kernel-based object tracking. *IEEE Trans. on Pattern Analysis and Machine Intelligence 25*, 5 (2003), 564–577.

22. Yang, C., Duraiswami, R., and Davis, L. Efficient mean-shift tracking via a new similarity measure. In *International Conference on Computer Vision and Pattern Recognition*, 2005, 176–183.

23. Pérez,P., Hue, C., Vermaak, J., and Gangnet, M. Color-based probabilistic tracking. In *European Conference on Computer Vision*, 2002, 661–675.

24. Nummiaro, K., Koller-Meier, E., and Gool, L.V. An adaptive color-based particle filter. *Image and Vision Computing 21*, 1 (2003), 99–110.

25. Liu, G., Tang, X., Cheng, H.D., Huang, J., and Liu, J. A novel approach for tracking high speed skaters in sports using a panning camera. *Pattern Recognition 42*, 11 (2009), 2922–2935.

26. Watanabe, T., Haseyama, M., and Kitajima, H. A soccer field tracking method with wire frame model from TV images. In *Proceedings IEEE International Conference on Image Processing*, 2004, 1633-1636.

27. Farin, D., Krabbe, S., de With, P.H., and Effelsberg, W. Robust camera calibration for sport videos using court models. In *SPIE Storage and Retrieval Methods and Applications for Multimedia*, 2004, 80–91.

28. Hayet, J.B., Piater, J., and Verly, J. Robust incremental rectification of sport video sequences. In *British Machine Vision Conference*, 2004, 687-696.

29. Shi, J. and Tomasi, C. Good features to track. In *International Conference on Computer Vision and Pattern Recognition*, 1994, 593-600.

30. Tomasi, C. and Kanade, T. Detection and tracking of point features. Technical Report CMU-CS-91-132, Carnegie Mellon University, 1991.

31. Fischler, M.A. and Bolles, R.C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. of the ACM 24*, 6 (1981), 381–395.

32. Hartley, R. and Zisserman, A. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

33. Harris, C. and Stephens, M. A combined corner and edge detector. In *Fourth Alvey Vision Conference*, 1988, 147–151.

34. Yilmaz, A., Javed, O., and and Shah, M. Object tracking: A survey. *ACM Comput. Surv. 38*, 4 (2006), 1-45.

35. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Journal of Computer Vision 60*, 2 (2004), 91–110.

36. Sivic, J. and Zisserman, A. Video google: A text retrieval approach to object matching in videos. In *International Conference on Computer Vision*, 2003, 1470–1477.

37. Brown, M. and Lowe, D.G. Recognising panoramas. In *International Conference on Computer Vision*, 2003, 1218–1225.

38. (http://www.isu.org/).

39. Bouwmans, T., El Baf, F., and Vachon, B. Background modeling using mixture of Gaussians for foreground detection - A survey. *Recent Patents on Computer Science 1*, 3 (2008), 219-237.

40. Wang, Y., Cheng, H.D., and Shan, J. Detecting shadows of moving vehicles based on HMM. In *International Conference on Pattern Recognition*, 2008, 1–4.

41. MacQueen, J. B. Some methods for classication and analysis of multivariate observations. In *Proceedings of the Fifth Symposium on Math, Statistics, and Probability*, 1967, 281-297.

42. Julier, S.J. and Uhlmann, J.K. Unscented filtering and nonlinear estimation. *Proceedings IEEE 92*, 3 (2004), 401-422.

43. Dunn, J.C. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics 3*, 3 (1973), 32-57.

44. Bezdek, J. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Kluwer, 1981.

45. Chan, T. and Vese, L. Active contours without edges. *IEEE Trans. Image Processing 10*, 2 (2001), 266-277.

46. Li, C., Kao, C., Gore, J. C. and Ding, Z. Minimization of region-scalable fitting energy for image segmentation. *IEEE Trans. Image Processing, 17* (2008), 1940–1949.

CURRICULUM VITAE

Yuxuan Wang
(2012)

EDUCATION

Ph.D. Computer Science May 2012(expected)
Utah State University (USU), Logan, UT GPA: 4.0
Dissertation: "*A Novel and Effective Short Track Speed Skating Tracking System*"
Advisor: Heng-Da Cheng

M.S. Computer Science July 2006
Harbin Institute of Technology (HIT), Harbin, China
Thesis: "*Discriminative Training of Hidden Markov Model in Chinese Sign Language Recognition*"

B.S. Computer Science July 2004
Harbin Institute of Technology (HIT), Harbin, China
Thesis: "*Protein Motif Recognition and Generation*"
GPA ranked top 10 among 400 graduates

RESEARCH INTERESTS

- Image Processing
- Computer Vision
- Machine Learning
- Bioinformatics

RESEARCH EXPERIENCE

Graduate Student/Research Assistant Sep 2006-present
Dr. Cheng's CVPRIP Research Group, Department of Computer Science, USU

- Designed and developed
  - An automatic vehicle detection and classification system. Cooperated with Utah Department of Transportation (UDOT), a video based vehicle classification system is developed, aiming at discovering large trucks. IDE is VS 2008.
  - A pavement detection system. Cooperated with UDOT, the system integrates line scan camera, GPS, encoder, splitter, and other equipment. Algorithms are carefully designed to handle complex road situations. Cracks are detected and classified automatically. IDE is VS 2008 and the SDK Sapera LT is employed for line scan camera developing.
  - A short track speed skating tracking system. Based on single panning camera, skaters in short track speed skating competitions are tracked automatically and lots of invaluable information can be generated such as trajectories, velocity analysis, and 2D reconstruction animations. The proposed novel algorithm

solve bunch of challenging problems such as skating in groups, wearing the same uniform, difficulty for registration, etc.. Currently, Matlab is employed for research purpose. Latter, C++will be employed to build the practical system.

- Participated in the development of
  - A novel and automatic segmentation algorithm for breast ultrasound images, using neutrosophy solving the fuzzy problems.
  - A completely automatic segmentation algorithm, using multiple domain features.
  - An automatic and efficient ROI (region of interest) generation method for breast ultrasound images using region growing.
  - A novel and robust seed point selection algorithm utilizing both texture and spatial features of breast ultrasound images.

Graduate Research Assistant Sep 2008-present

Dr. Brett's IDIAS Research Group, Department of Instructional Technology and Learning Science, USU

- Lead or Participate in the development of
  - A 3D game Engine "HEAT". Focus on developing VOIP system to allow communicating with each other though talkback system. The engine is based on C++ and many 3rd party open sources.
  - A recommendation system for open education resources. Focus on new algorithm design to improve recommendation results. The web service is based on Ruby on Rails, and JAVA is employed for algorithm designs.
  - A mobile app developing project called Geobob. GeoBob is a development platform that enables the design and development of location-based applications for a wide variety of uses on mobile devices like iPhone and Android phones. I was focusing on iPhone side developing. As a by-product, the official USU app was released in early 2011 and got thousands of downloads. Another one GeePerS Math, which is a GPS based orientation game, will be released soon.

Dr. Yan's Bioinformatics Research Group, Department of Computer Science, USU, 2007

- Designed and developed:
  - A system to analyze the protein-DNA interaction using structure and physicochemical features of proteins, using support vector machine and decision tree, respectively. According to the study, four features were finally selected to be good features which can predict the specificity of amino acid-base contacts.

Graduate Research Assistant September 2004 – July 2006

VILAB, Department of Computer Science, HIT, Harbin, China

- Focus on Chinese sign language recognition project. Data gloves are employed to obtain spatial-temporal sign language data. All 5000 Chinese sign language words are trained and tested. A novel discriminative HMM training method is developed

to improve the recognition rate. The system is based on windows and the IDE is VC++ 6.0.

Undergraduate Research Assistant September 2003 – July 2004

VILAB, Department of Computer Science, HIT, Harbin, China

- Developed a system to discover motif of protein structures. The system is based on windows and the IDE is VC++ 6.0. The thesis won the "Outstanding Bachelor Thesis of year" award.

TEACHING EXPERIENCE

Teaching Assistant

Department of Computer Science, USU

Collaborated on class development and group discussion, met with students upon request, and graded all assignments and programming projects for:

- JAVA/Internet Programming (CS3410)
- Advanced Bioinformatics (CS6670)
- Computer Vision, Pattern Recognition and Image Processing (CS5650)
- Algorithms and Data Structure (CS2420)
- Fuzzy Logic (CS6630)
- Advanced Algorithms (CS5050)

USU International Teaching Assistant Workshop Fall 2008

Prepared syllabus, lecture slides, and gave sample lectures on a chosen course.

Got TA Certificate.

PUBLICATIONS

Journals

1. **Y. X. Wang**, H. D. Cheng and J. Shan, "A Novel and Effective Short Track Speed Skating Tracking System", *submitted to Pattern Recognition*.
2. J. Shan, H. D. Cheng and **Y. X. Wang**, "A Novel Segmentation Method for Breast Ultrasound Images Based on Neutrosophic L-Means Clustering", *Medical Physics, revision*.
3. J. Shan, H. D. Cheng and **Y. X. Wang**, "Completely Automated Segmentation Approach for Breast Ultrasound Images Using Multiple-Domain Features", *Ultrasound in Medicine and Biology,* Volume 38, Issue 2, Pages 262-275, February 2012.

Conferences

1. J. Shan, **Y. X. Wang** and H. D. Cheng, "Completely Automatic Segmentation for Breast Ultrasound Using Multiple-Domain Features", *Int. Conf. on Image Processing (ICIP2010)*, Hong Kong, China, September 26-29, 2010. (*Oral presentation, acceptance rate 47%*)

2. Brett E. Shelton, Joel Duffin, **Y. X. Wang**, Justin Ball: Linking open course wares and open education resources: creating an effective search and recommendation system. *Procedia CS 1*(2): 2865-2870, 2010.
3. **Y.X. Wang**, H. D. Cheng and J. Shan, "Detecting shadows of moving vehicles based on HMM", *Proc. of the 19th International Conference on Pattern Recognition (ICPR 2008)*, Florida, USA, Dec 8-11, 2008. (*Acceptance rate 38%*)
4. J. Shan, H. D. Cheng and **Y.X. Wang**, "A novel automatic seed point selection algorithm for breast ultrasound images", *Proc. of the 19th International Conference on Pattern Recognition (ICPR 2008)*, Florida, USA, Dec 8-11, 2008. (*Acceptance rate 38%*)
5. J. Shan, H. D. Cheng and **Y.X. Wang**, "A completely automatic segmentation method for breast ultrasound images using region growing", *11th Joint Conference on Information Science,* Shenzhen, China, Dec 15-20, 2008.
6. J. Shan, **Y.X. Wang** and C. Yan, "Toward the recognition code of protein-DNA recognition", *Proc. of IEEE 7th International Symposium on BioInformatics and BioEngineering*, pp. 1290 – 1293, 2007. (*Acceptance rate 13%*)
7. **Y.X. Wang**, H. D. Cheng and Chris Glazier, "An Improved Approach for Vehicle Detection and Classification", *Joint Conference on Information Science*, 2007.

AWARDS AND HONORS

- **Outstanding Graduation** of Heilongjiang Province, China, 2004
- **Outstanding Bachelor Thesis of year** of Heilongjiang Province, China, 2004
- **People Fellowship, 8 times,** Harbin Institute of Technology, China, 2000-2004
- **Guang-Hua Fellowship**, Harbin Institute of Technology, China, 2002-2003
- **Hua-Wei Endowed Scholarship,** Harbin Institute of Technology, China, 2001-2002
- **Excellent Student Activity Leader**, Harbin Institute of Technology, China, 2002-2004
- **Won the Top10 Famous Class Prize as class monitor,** Harbin Institute of Technology, China, 2000-2001

PROGRAMMING SKILLS

Proficient: Matlab, C/C++, Objective C
Familiar: Java, R, C#, SQL, Ruby on Rails
Platform: Windows, Linux systems, Mac/IOS

AD HOC REVIEWER

International Conference on Computer Vision, Pattern Recognition & Image Processing 2007
International Journal of Pattern Recognition and Artificial Intelligence
Pattern Recognition

SERVICE AND ACTIVITIES

- Monitor of the 7th class of Computer Science Department, Harbin Institute of Technology, China, 2000-2004
- Officer of Student Union of Computer Science Department, Harbin Institute of Technology, China, 2001-2002

LANGUAGES

English (fluent), Chinese (native)