Elektrotehniški vestnik 67(5): 281–288, 2000
Electrotechnical Review, Ljubljana, Slovenija

A System for Tracking Players in Sports Games by Computer Vision

Janez Perš, Stanislav Kovačič

University of Ljubljana, Faculty of Electrical Engineering,Tržaška 25, 1000 Ljubljana, Slovenija

Abstract. We present a computer vision system for tracking players in indoor team games, e.g. handball in our particular case. We describe several image processing and tracking methods, along with camera calibration and lens distortion correction. The output of the system consists of spatio-temporal trajectories of the players, which can be further processed and analyzed by sport experts. We present some experimental results as well.

Key words: computer vision, object tracking, human motion analysis, camera calibration, handball

Sistem za sledenje igralcev v športnih igrah s pomočjo računalniškega vida

Povzetek. Predstavljen je razvoj sistema računalniškega vida za sledenje igralcev v ekipnih dvoranskih športih, še posebej v rokometu. Opisanih je nekaj uporabljenih metod sledenja in obdelave slik, skupaj s kalibracijo kamere in korekcijo radialnega popačenja. Izhod sistema predstavljajo trajektorije igralcev v koordinatnem sistemu igrišča. Trajektorije so športnim strokovnjakom osnova za nadaljnjo analizo poteka tekme. Prikazani so nekateri eksperimentalni rezultati.

Ključne besede: računalniški vid, sledenje objektov, analiza človeškega gibanja, kalibracija kamere, rokomet

1 Introduction

The human motion analysis is receiving increasing attention from computer vision researchers. This interest is motivated by applications over a wide spectrum of topics [1, 2], e.g., motion analysis involving human body parts, tracking of human motion, and recognizing human activities from image sequences. Application of the computer vision technology to sports domain presents a useful and challenging area of research for several reasons. Most sports involve complex human motion and therefore automated capturing, analyzing and quantifying the ability of the athlete can offer significant help to the sports expert. In some sports, most notably in athletic disciplines, the goal is to develop models of the human body that explain how it functions mechanically and potentially to increase its movement efficiency [3]. This involves segmenting the body parts, tracking the joints in the sequence of images and recovering the 3D body structure. Whole-body tracking and tracking moving individuals without considering the geometric structure of the body, is particularly useful in analysis of team sports, where the positions of players as a functions of time have to be determined. If the trajectories are sufficiently accurate, a wealth of additional information, e.g., players velocity, acceleration and players interactions, can be obtained. The trajectories are also useful in automatic highlighting and annotation of sport events, where the aim is to help the audience to better appreciate the sport [4, 5]. Hence, the accuracy becomes less important in this case.

We concentrate our activities on tracking people in team sports based on computer vision. For many years the analysis of a sport event has been based on "observation sheets" filled-in during the match. In the 1980's, modern techniques of the motion analysis were developed with the help of video recordings [6, 7]. Motion acquisition and analysis were performed manually, which was a time consuming and tedious task. In the past, progress in introducing the computer vision technology to the team sports domain was slow, mainly due to inadequate video and computational facilities, as just a single match may require processing of tens of thousands of complex images [8]. Large amounts of data and high computational load are by no means the only burden. The players strive to move rapidly, change direction unpredictably and collide with one another. They violate the smooth motion assumption, on which

Received 4 Jun 2000
Accepted 20 October 2000

many tracking algorithms are based. Players appear in the images as highly non-rigid forms, especially due to the movements of their extremities. In addition, cameras used to record the sport events have to cover a large area, either by following the players of interest, or by using wide-angle lenses, which results in substantial image distortion and low resolution. Many of the proposed approaches solved the motion acquisition and analysis problem only partially and were therefore unable to provide an adequate solution to the sports experts, i.e. tracking *every* player and/or the ball in the *whole* field, and in *every* instance of time [6]. At least one commercial solution based on computer vision for obtaining and analyzing trajectory data in soccer exists (AMISCO [9]), but little is known about the degree of human intervention required to maintain error-free tracking. Solutions, which require players to wear special transmitters (for example TRAKUS [10]), are inappropriate for use in regular matches.

In this paper we focus on tracking the players in the handball game. Advances in the computer and video technology over the past several years enabled us to approach the problem by digitizing the whole handball match at 25 frames per second, using two wide-angle cameras. Acquired data was subsequently used to develop a set of tracking methods, which suit best the nature of the game. The assumption underlying our approach is that it is possible to infer 2-D motion in the court plane of otherwise 3-D objects from 2-D images.

The rest of the paper is organized as follows: first, we define the requirements of the tracking system for tracking players in handball. We discuss the camera setup and camera calibration. Next, we describe three different algorithms for player tracking and finally, we present the evaluation of these algorithms.

2 Problem description

Our long-term goal is to develop a computer vision system, which will be ultimately capable of tracking each and all players with known accuracy for the duration of the whole match, and with minor operator intervention, possibly in real time. The output result of the tracking system will be spatio-temporal trajectories for all players, which could be further analyzed in the context of some particular sport, i.e. handball in our case. These player trajectories provide the sports expert with valuable data of the player capabilities and performance. The output results should satisfy several requirements. The trajectories obtained by tracking should be accurate enough to enable a simple movement analysis (velocity and acceleration calculation) and highly reliable - no swapping of players is allowed. The non-intrusive method of tracking players during the match is also preferred, since it can be used in regular matches where the rules strictly prohibit any kind of intrusive measurement.

Given the requirements, specified above, the main objectives in development of the tracking system are:

- To choose the camera setup, which would not additionally complicate the tracking problem and would guarantee reasonably reliable results.
- To develop techniques for camera calibration and lens distortion correction, as very wide angle lenses are used.
- To develop a set of off-line tracking methods, which would suit the specified requirements.

2.1 The game

European or team handball is an indoor sport, in which two teams compete against each other. The playing court dimensions are 20 by 40 meters, as shown in Fig. 1a. A team consists of 12 players. No more than seven players (6 court players and 1 goalkeeper) of each team may be present on the court at the same time. The playing time is 2 halves of 30 minutes with a half-time break of 10 minutes.

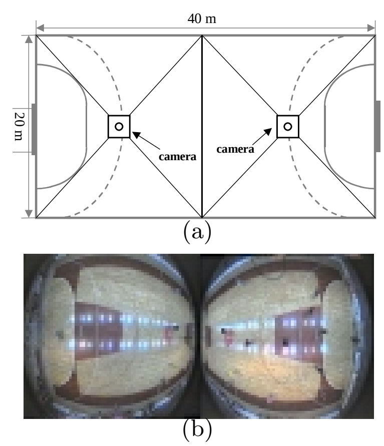

Figure 1. (a) Handball playing court and camera placement. (b) Example of combined image from two cameras, taken at the same instant of time.

3 Image acquisition

Proper image acquisition significantly influences the performance of the tracking algorithms. In case of video annotation and highlighting [5] high accuracy is not required. In case of player motion acquisition and analysis, however, where certain measurements are performed and degree of uncertainty has to be specified, careful planning of image acquisition proves to be crucial for the success of the whole system [12].

3.1 Camera setup

Different camera setups are possible to record the match. To determine the trajectories of the players (or the ball) from the beginning to the end of the game, the objects of interest have to be in the field-of-view for the duration of the whole match. This can be achieved either by camera movement [4] or by use of wide-angle lens [6]. Camera movement can add significant degree of difficulty to the tracking problem, as the objects and the sensor are independently moving with respect to the reference coordinate frame. Therefore, continuous calibration is required. Moreover, some motion detection techniques (for example, subtracting the adjacent frames [11]) become nearly impossible to implement. To alleviate these problems, two stationary cameras with wide-angle lenses were chosen.

Camera angle with respect to the court is another important question. For practical reasons some researchers have placed the cameras at the angle smaller than 90 degrees [6, 4], although placement directly above the court with optical axis perpendicular to the court plane would be optimal for player tracking. This resulted in ill-conditioned tracking, as reported in [12]. Therefore, we placed the cameras directly above the court, as shown in Fig. 1. Cameras were equipped with 103° lenses, because they could be mounted no higher than 10 meters from the court plane.

3.2 Video recording and capture

The whole handball match that lasted for about an hour was recorded using two PAL cameras and two S-VHS videorecorders. A transfer to digital domain was carried out using the Motion-JPEG video acquisition hardware, at 50 fields per second and image resolution of 384x288 pixel. The combined image from both cameras is shown in Fig. 1b.

4 Camera calibration

To perform measurements based on the acquired images, the relations between pixel coordinates in each of the images and world (court) coordinates have to be known. These relations are obtained by the camera calibration. The procedure is simplified due to rigid sport regulations, which precisely specify the locations and dimensions of various landmarks on the playing court. Unfortunately, due to the large radial distortion otherwise widely used calibration technique [13] fails to produce satisfactory results. Although some authors [14] suggested an extension to the popular method, we decided to take a different approach by modeling the radial image distortion more accurately. To determine the scaling factor, camera orientation and position with respect to the court plane, we employed a linear camera calibration model.

Fig. 2 illustrates the problem of radial distortion. Let us imagine an ideal pinhole camera, mounted on a pan-tilt device. Point 0 is the point of intersection of optical axis of the camera with the court plane, when the pan-tilt device is in its vertical position. Point C denotes the location of the camera, and X is the observed point on the court plane, at distance R from the point 0. H is the distance from the camera to the court plane. Angle α is the angle of the pan-tilt device when observing the point X. The differential dR of radius R is projected to the differential dr, which is parallel to the camera image plane. The image of dr appears on the image plane. Relations between dR, dr and α are given within the triangle on the enlarged part of Fig. 2.

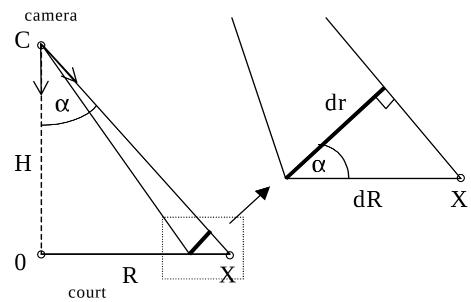

Figure 2. A model of radial distortion.

Thus, we can write the following relations:

$$dr = cos(\alpha) \cdot dR, \quad \alpha = arctg(\frac{R}{H}), \tag{1}$$

$$dr = cos(arctg(\frac{R}{H}))dR. \tag{2}$$

Let us substitute the pan-tilt camera with a fixed camera, equipped with wide-angle lens. The whole area, which is covered by changing the angle α of the pan-tilt camera, is captured simultaneously to the single image of the stationary camera. Additionally, let us assume that the scaling factor between the dr and the image of dr on the image plane equals 1. Therefore, we can obtain the length of the image of radius R on the image plane by integrating the left side of Eq. (2) over the interval $(0, r_1)$, and the right side over the interval $(0, R_1)$,

$$\int_0^{r1} dr = \int_0^{R1} cos(arctg(\frac{R}{H}))dR. \tag{3}$$

With R_1 being the distance from the observed point X to the point 0 and r_1 being the distance from the image of point X to the image of point 0 on the image plane, we have the solution of the inverse problem:

$$r_1 = H \cdot ln \left(\frac{R_1}{H} + \sqrt{1 + \frac{R_1^2}{H^2}} \right). \qquad (4)$$

By solving Eq. (4) for R_1 we obtain the formula, which can be used to correct the radial distortion:

$$R_1 = \frac{H}{2} \frac{(e^{-\frac{2r_1}{H}}) - 1}{e^{-\frac{r_1}{H}}}. \qquad (5)$$

For illustrative purposes, a result of radial distortion correction is shown in Fig. 3. Nevertheless, we decided to perform tracking on raw, uncorrected images, and to correct the obtained player positions thereafter.

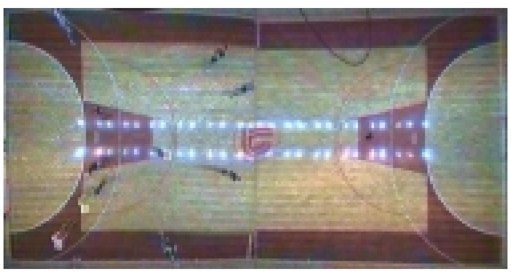

Figure 3. A combined image from both cameras after the radial distortion correction.

5 Player tracking

Development and evaluation of various tracking techniques was in the focus of research concerning the use of computer vision techniques in the human motion analysis [4, 14, 15]. Various general purpose tracking algorithms exist [16], giving the researchers the opportunity to choose from a wide range of methods. We decided to combine three different algorithms and used them in player tracker: motion detection, template tracking and color-based tracking.

5.1 Motion detection

The most straightforward approach to motion detection is by subtracting adjacent frames in the image sequence [11]. This approach is a general one, fast and easy to implement, but the reliability in case of slowly moving objects is rather low. Therefore, a modified approach, thoroughly explained in [8] was used. Each frame in the sequence was subtracted from the "reference frame", i.e. an image of the empty court, yielding the difference image. The difference image was then thresholded, filtered (which is a time consuming operation) and the resulting blobs, which correspond to players, were analyzed. Unfortunately, this movement detection and tracking

Figure 4. A typical scene which confuses the motion detection algorithm. a) Four players in an image are marked by white arrows. b) The difference image for this situation. Series of light reflections, seen in lower-left quarter of the image render the situation even more difficult.

technique is not error-free. Collisions of more than two players and intense player shadows contacting other players require almost always an intervention by a human operator. Sudden leaps in player position occur, when the tracker confuses two players or the player and some other object (most often shadow or light reflection). A typical situation is shown in Fig. 4.

5.2 Template tracking

Most of the errors in the motion detection phase of the tracker occur when the tracker confuses players with other, visually different "objects" in image. This is not surprising, considering the non-discriminatory nature of the motion detection algorithm. Therefore, visual differences between the players and the background are exploited to further improve the tracking process.

Feature set, which can be used to successfully separate objects from the background, needs to be found. Due to low resolution and rapidly changing appearance of the players it is extremely difficult to build an accurate model of a handball player. Instead, we have experimentally defined a set of 2D functions, i.e. "templates", shown in Fig. 5, which extract the very basic appearances of the players. They resemble the Walsh functions, but they are not orthogonal.

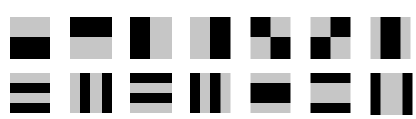

Figure 5. Basic templates of the player. Black areas represent zeros, while gray areas denote value of 1.

First, the region of interest (ROI) which surrounds the position of the player is defined.

Considering the size of the players in captured image, the region size was set to 16x16 pixels, with player position in the center of the region. Each channel of the RGB color image is processed separately and the vector $\boldsymbol{F}$, consisting of 14 features for each channel, is obtained using the following formula:

$$F_{i+14j} = \sum_{x=1}^{16} \sum_{y=1}^{16} K_i(x, y) \cdot I_j(x, y), \qquad (6)$$

where K_i is one of the 14 functions $(i = 0 \ldots 13)$, and I_j is one of the three RGB channels $(j = 0, 1, 2)$, obtained with respect to ROI from the current image. Each channel yields 14 features, which results in 42-dimensional feature vector $\boldsymbol{F}$.

Let vector $\boldsymbol{H}$ represent the estimated appearance of the player, and vector $\boldsymbol{G}$ represent the appearance of the background (playing court) at the same coordinates. Our goal is to classify the unknown object in the region of interest I, represented as vector $\boldsymbol{F}$, either as a “player” or a “background”. The simplified, two-dimensional case is shown in Fig. 6.

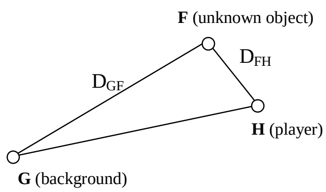

Figure 6. Classification of an unknown object.

Vector of features $\boldsymbol{G}$ is calculated from the reference image of the playing court at the same coordinates as $\boldsymbol{F}$. The reference vector $\boldsymbol{H}$ is obtained by averaging the last n vectors of features for a successfully located player, which allows certain adaptivity, as the player appearance changes over time. The value of n depends on the velocity at which the players move, but best results were obtained with the value of $n = 50$ which corresponds to two seconds of video sequence.

Similarity measure S is obtained using the following formula:

$$S = \frac{D_{FH}}{D_{GF} + D_{FH}}, \qquad S \in [0, 1], \qquad (7)$$

where D_{GF} and D_{FH} are Euclidean distances. The domain of measure S is interval $[0 \ldots 1]$. Low value of S means high similarity between the observed object $\boldsymbol{F}$ from the region of interest I and stored appearance of the player in the vector $\boldsymbol{H}$. Fig. 7 shows the test result on a single player and demonstrates the ability of this technique to locate an object.

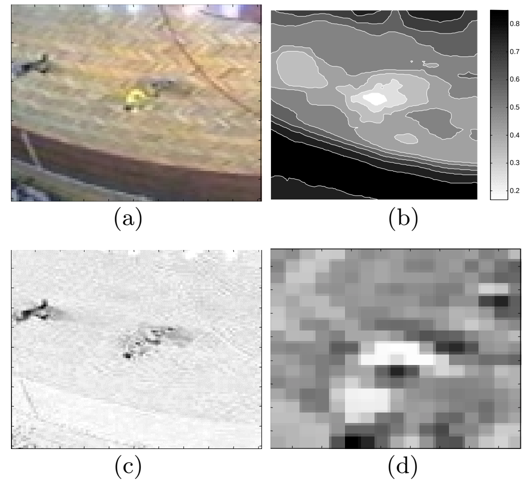

Figure 7. Locating the player wearing a yellow dress using the set of templates. (a) The player is shown in the center of the image. (b) Similarity measure S for the template tracking method, as defined in (7). Feature vector $\boldsymbol{H}$ (player reference) was obtained from the subimage of the same player, captured at different instants of time. The white area corresponds to the region of high similarity (S is 0.2 or lower) and marks the player location more accurately than the difference image. (c) The difference image for that particular case. (d) Distance to the yellow color for a close-up view of a player from the image (a). White pixels mark the areas that match the yellow color, while darker ones mark areas of different colors. The image was obtained using the similarity measure (8)

Starting with the previous player position, the 3x3 neighborhood is examined for the minimum of similarity measure (7). The position of the minimum is used as the next estimate and the process is iteratively repeated up to 10 times. The maximum number of iterations limits the area that is being searched and is defined by the maximum expected player movement.

The main advantage of this method is an accurate position measurement and a low probability of sudden leaps in the player position. Its main weakness is the possibility of overadaptation to the background (the distance between $\boldsymbol{H}$ and $\boldsymbol{G}$ becomes too small, (see Eq. (7) and Fig. 6), which can cause gradual drift in player position.

5.3 Color tracking

Color as an identifying feature [17] can be also used for the task of tracking players. Color is generally largely independent of the view and resolution, and remains constant over long intervals of time. Therefore, colors of the players dresses could be input to the tracking system at the beginning of the tracking process without any adaptations later.

Color identification and localization, based on color histograms, was reported by Swain and Ballard [17]. However, given a small number of pixels that comprise each of the players, this technique is not appropriate. In most cases, there are only a few (3-6) pixels that closely resemble the reference color of the player's dress. The situation is illustrated by Fig. 7d. The 9x9 neighborhood of the previous player position is examined and only the location of the pixel that is most similar to the player's color becomes a new player's position. The three-dimensional RGB color representation was chosen instead of HSI, as some players wear dark dresses, which would result in undefined values of hue. The similarity measure is defined as:

$$S(x,y) = ((I_R(x,y) - C_R)^2 + \\ +(I_G(x,y) - C_G)^2 + (I_B(x,y) - C_B)^2)^{\frac{1}{2}}, \tag{8}$$

where I is the image and C is the color of the player. R, G and B denote the red, green and blue channel, respectively.

The algorithm searches for the pixel most similar to the recorded color of the player. The search is performed in a very limited area around the previous player position.

The advantage of described algorithm is high reliability. The algorithm tracks players successfully even when the apparent player color is changed due to signal distortion during tape recording or lossy compression. The main problem is caused by diverse background with colored areas, which closely correspond to the color of the player's dress. The disadvantage of this method is also a high amount of jitter in the resulting player trajectories.

5.4 Trajectory post-processing

The trajectories obtained with either of the previously described methods contain high amount of noise. The noise level depends on the actual method or combination of methods used, but without a proper filtering, none of the trajectories can be used for velocity calculation. The distance, traveled by the player during a certain period is also affected, as the distance is calculated by integrating the magnitude of velocity.

The spectrum of noise overlaps with the spectrum of rapid player movements and some data loss is expected when filtering player trajectories. An obvious way of trajectory filtering is by using a Gaussian filter, as shown in (9). We process x and y components of the trajectory separately, treating them as one-dimensional, time-dependent signals.

$$u(t) = \frac{1}{2N_F + 1} \sum_{i=-N_F}^{N_F} x(t+i) \cdot G(i), \tag{9}$$

where N_F denotes the width of the filter, u is the filtered component of the trajectory, x is the component of the raw trajectory as provided by the tracking method and G is the array of Gaussian coefficients. The precalculated set of $N_F + 1$ coefficients in the range of Gaussian function $(-3\sigma, 3\sigma)$ was used. A wider filter yields more stable results in terms of velocity and distance calculation. However, as rapid player movements overlap with noise, a wider filter also eliminates the presence of these important elements of the player movement.

6 Results

To evaluate the performance of our tracker, three experiments were performed.

Two video sequences, extracted from the recording of the whole handball match, were used in the testing. Duration of the sequence 1 was 30 seconds, and duration of the sequence 2 was 50 seconds. Both were sampled at 25 frames per second. Some players from one team were wearing differently colored dresses. All other players were wearing the same dark blue dresses. In the first sequence, tracking was done simultaneously for all 14 players of both teams. Operator interventions required to maintain error-free tracking process were counted to evaluate tracking reliability. In the second experiment, automatic tracking with operator interventions disabled was performed and the processing time per frame was measured. The computational efficiency tests were done on the PC with Pentium III running at 500 MHz. The tracking program was written and compiled under Borland Delphi 3 with optimization enabled. The results are shown in Table 1, where capital letters denote the following tracking methods:

- A: Tracking using the motion detection alone.
- B: Color tracking alone.
- C: Combination of color and template tracking.

Method C combines position estimation using color tracking and fine position measurement using the template tracking method. However, the initial point for determining the position in the next frame is the output of the color tracking algorithm alone, to eliminate drift in position.

To further test the performance of our tracker, we compared the computer-generated trajectories to manually entered positions using a mouse as a pointing device. Five human operators performed manual tracking of one of the players in the sequence 2. The sampling rate for the manual tracking was set to 2 frames per second to keep the number of required mouse clicks reasonably low. The automatic tracking using all three methods was performed as well. Only positions of the player in those frames that corresponded to frames shown to operators during manual tracking were extracted to ensure unbiased comparison. The results obtained by the operators were used to calculate the "mean path" of the player, which was used as a true-position estimate:

$$\overline{x}(k) = \frac{1}{5}\sum_{i=1}^{5} x_i(k), \quad \overline{y}(k) = \frac{1}{5}\sum_{i=1}^{5} y_i(k), \tag{10}$$

where $k = 1 \ldots k_{max}$ and k_{max} denotes the total number of frames used in the calculation. Next, the mean position error was calculated for five operators and three automatic methods:

$$E_n = \frac{1}{k_{max}} \sum_{k=1}^{k_{max}} \sqrt{(x(k) - \overline{x}(k))^2 + (y(k) - \overline{y}(k))^2}, \tag{11}$$

where $n = 1 \ldots 8$. As the jitter in computer generated output trajectories increases the length of player path, we decided to compare the lengths of paths obtained by different human operators and lengths of unfiltered paths of the three automatic methods. The path length was calculated as follows:

$$L = \sum_{k=2}^{k_{max}} \sqrt{(x(k) - x(k-1))^2 + (y(k) - y(k-1))^2}. \tag{12}$$

The mean error and the path length for each human operator and each of the methods is shown in Table 1.

Method	No. of interventions	Efficiency [sec/frame]	Mean pos. error [cm]	Path length [m]
A	45	0.424	35.3	73
B	12	0.175	33.4	82
C	14	0.229	28.0	75
O1	X	X	9.8	76
O2	X	X	9.0	78
O3	X	X	7.3	76
O4	X	X	8.1	77
O5	X	X	6.7	77

Table 1. Comparison results for three different tracking methods (A-C) and five human operators (O1-O5).

The advantages and disadvantages for each of the methods can be clearly seen. Method A gives stable results (the shortest path), as the center-of-gravity approach introduces low-pass filtering of the output data. On the other hand, in case of tracking all 14 players, more than one intervention per second of playing time is required, which puts substantial pressure on the human operator supervising the tracking process. The method A is the slowest, mainly due to the time consuming iterative image filtering operations.

The color tracking method (B) required a small number of interventions, as expected. It is also fast. On the other hand, the results provided by this method contain a higher amount of jitter and require extensive filtering.

The method C is nearly as fast as B. It contains a lower amount of jitter and requires little human operator intervention.

Based on these observations, we conclude that the combination C is the most suitable for use in the automated player tracking.

7 Conclusions

Development of an automated tracking system for tracking players in sports games is presented. It forms a core of the complete sports analysis system, which will include advanced data processing, statistical analysis, and presentation capabilities. Although the system was only tested in the handball domain, we expect that it can be equally useful in other indoor sports, like basketball, indoor hockey and indoor football. If cameras can be placed directly above the playing court, the system can be used in outdoor sports, too.

8 Acknowledgements

A close collaboration with the sports experts was extremely important during the development process. They contributed valuable suggestions and a continuos feedback about the system. We thank Marta Bon and Marko Šibila from the Faculty of Sports at the University of Ljubljana for their valuable contribution.

9 References

[1] J. K. Aggarval, Q. Cai, Human motion analysis: A review, *IEEE Nonrigid and Articulated Motion Workshop,* pages 90-102, Puerto Rico, June 17-19 1997.

[2] Q. Cai, J.K Aggarwal, Tracking human motion using multiple cameras, *Proceedings of the 13th International Conference on Pattern Recognition (ICPR'96),* pages 68-72,Vienna.

[3] T. W. Calvert, A. E. Chapman: Analysis and synthesis of human movement, In*Handbook of Pattern recognition and Image Processing: Computer Vision,* Edited by T. Y. Young, Ch. 12, pages 432-474, 1994.

[4] S. S. Intille, A. F. Bobick, Visual tracking using closed-worlds, *Proc. of the Fifth International Conference on Computer Vision*, MIT, Cambridge, MA, pp. 672-678, June 20-23, 1995.

[5] G. S. Pingali, Y. Jean, I. Carlborn, Real time tracking for enhanced tennis broadcasts, *Proceedings of CVPR 98: 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* pages 260-265, 1998.

[6] W. S. Erdmann, Gathering of kinematic data of sport event by televising the whole pitch and track, *Proc. of 10th ISBS symposium*, pages 159-162, Rome, 1992.

[7] A. Ali, M. Farrally, A computer-video aided time motion analysis technique for match analysis, *The Journal of Sports Medicine and Physical Fitness,* Vol. 31, No. 1, March 1991, pages 82-88.

[8] J. Perš, S. Kovačič, Computer vision system for tracking players in sports games, *Proc. of ERK '99,* , pages 253-256, 1999

[9] J. Razinger, Nogomet z računalniško podporo, *Življenje in Tehnika, št. 7-8/1998, s*trani 40-43.

[10] Trakus, Inc. http://www.trakus.com.

[11] J. Perš, S. Kovačič, Sledenje objektov z metodami umetnega vida, *Proc. of ERK '98*, pages 469-470, 1998.

[12] MIT Media Laboratory, Computers watching football - the data, http://vismod.www.media.mit.edu/vismod/-demos/football/manual.htm

[13] R. Y. Tsai, A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses, *IEEE Journal of Robotics and Automation,* Vol. RA-3, No. 4, 1987, pages 323-344.

[14] K. Grešak, Analiza športnih iger z računalniškim vidom, diplomsko delo, *Fakulteta za računalnistvo in informatiko, Univerza v Ljubljani, 1999.*

[15] J. Segen, S. (Gopal) Pingali, A camera-based system for tracking people in real time, *Proceedings of 13th International Conference on Pattern Recognition (ICPR'96)*, pages 63-67, Vienna.

[16] W. N. Martin, J. K. Aggarwal, editors, Motion understanding: Robot and Human Vision, Boston, *Kluwer Academic Publishers,* 1998.

[17] M. J. Swain, D. H. Ballard, Color indexing, *International Journal of Computer Vision,* Vol. 7, No. 1, November 1991, pages 11-32.

Web links verified on March 15, 2000.