Real Time Colour Based Player Tracking in Indoor Sports

Catarina B. Santiago, Armando Sousa, Luís Paulo Reis, and Maria Luísa Estriga

Abstract In recent years there has been a growing interest by the sport's experts (teachers and coaches) in developing automatic systems for detecting, tracking and identifying player's movements with the purpose of improving the players' performance and accomplishing a consistent and standard analysis of the game metrics. A challenge like this requires sophisticated techniques from the areas of image processing and artificial intelligence. The objective of our work is to study and develop hardware and software techniques in order to build an automatic visual system for detecting and tracking players in indoor sports games that can aid coaches to analyse and improve the players' performance. Our methodology is based on colour features and therefore several colour image processing techniques such as background subtraction, blob colour definition (RGB and HSL colour spaces) and colour blob manipulation are employed in order to detect the players. Past information

C.B. Santiago (✉) and A. Sousa
FEUP – Faculty of Engineering of the University of Porto, Rua Dr. Roberto Frias, s/n 4200-465, Porto, Portugal
and
INESC – Institute for Systems and Computer Engineering of Porto, Campus da FEUP, Rua Dr. Roberto Frias, 378, 4200-465, Porto, Portugal
e-mail: catarina.santiago@fe.up.pt; asousa@fe.up.pt

L.P. Reis
FEUP – Faculty of Engineering of the University of Porto, Rua Dr. Roberto Frias, s/n 4200-465, Porto, Portugal
and
LIACC – Artificial Intelligence and Computer Science Lab, University of Porto, Portugal, Rua Dr. Roberto Frias, s/n 4200-465, Porto, Portugal
e-mail: lpreis@fe.up.pt

M.L. Estriga
FADEUP – Faculty of Sports of the University of Porto, Rua Dr. Plácido Costa, 91, 4200-450, Porto, Portugal
and
CIFI2D – Centre of Research, Education, Innovation and Intervention in Sport, Sports Faculty, University of Porto, Rua Dr. Plácido Costa, 91, 4200-450, Porto, Portugal
e-mail: lestriga@fade.up.pt

J.M.R.S. Tavares and R.M.N. Jorge (eds.), *Computational Vision and Medical Image Processing: Recent Trends*, Computational Methods in Applied Sciences 19,
DOI 10.1007/978-94-007-0011-6_2,

and players' velocity allow the tracking algorithm to define probable areas. Tests conducted with a single IP surveillance camera on the sports hall of the Faculty of Sports from the University of Porto showed detection rates from 72.2% to 93.3%.

Keywords Colour image processing · Object tracking · Game analysis

1 Introduction

In sports, especially at high level competition a small advantage of one team regarding another can be of great importance and decisive for a match or even a complete championship.

Therefore, teams are always trying to improve their performance and tactics in order to gain this advantage. So, it could be of greater help to record the game sequence from strategic points where the field can be completely surveyed and where the movements of all players may be recorded with precision. This way, after the game, the team's coach can identify their weak points and define measures to improve the team global behaviour.

Besides acquiring high quality images it would also be interesting to track the players and analyze their behaviour during the game: covered area, number of shoots and goals, number of passes, interaction between players, etc.

An automatic system such as this would bring many advantages, namely it would be able to handle a huge amount of data and perform a systematic evaluation, that is a very time consuming task and not always systematic when performed by a human being.

In this chapter we present the approach used for developing an automatic and intelligent visual system for detecting and tracking players in indoor sports games.

The main objectives are to design a vision system that is able to cover the entire field, detect each player individually and track his/her movements. One of the requirements for this system is that it must not cause any interference in the game, no special tags or colours should be placed neither in the area of action nor on the players, since most of the championships do not allow it.

A system like this is highly complex since it involves objects that are all very similar, constantly moving, changing of shape and getting together which makes an hard task to individualize and consistently identify and track each player. In fact, occlusion and player merging make the task of player identification very difficult.

This chapter is divided into eight sections. The initial sections introduce the topic under analysis including some background information, motivation and related work. The two subsequent sections (3 and 4) provide a description of the projected and implemented architectures and give a detailed explanation of the image processing system which includes the players' detection and tracking. Section 5 shows the results achieved so far and the last section refers to the conclusions and some future work.

2 Related Work

In recent years there has been a growing interest in developing automatic systems for detecting, tracking and identifying player's movements. Some works on the specific area of player tracking have been developed but the majority is intended to outdoor sports namely soccer. In outdoor sports it is very hard to mount a camera system and therefore most of the systems are developed upon images provided by broadcast TV cameras, either single camera or multiple camera systems.

Liu, Jiu et al. [1] use an approach based on a single moving broadcast camera system. Their system doesn't need to be manually initialized since it is able to learn the models of both the background and the players. The background is modelled using a dominant colour learning algorithm and a Haar feature based boosting cascade algorithm is able to represent the players. After this initial learning phase to model the players and the background, the detection and tracking are built upon a Markov Chain Monte Carlo data association method.

FIFA World Cup 2006 was their test platform and from the results their method seems to have high detection and labelling precision, around 90%.

On the opposite side Iwase and Saito [2] propose a dedicated eight camera system which is able to minimize the effects of occlusion. Each camera is treated as an independent system called inner-camera and is responsible for detecting and tracking the players based on very simple features such as colour, area and distance a player has moved. Whenever the inner-camera system is not able to detect the player due to occlusions, not detections or by the players being outside the angle of view an inter-camera process is, in most of the cases, able to identify the players. The geometrical relationships between the cameras is calculated based on planar homography in projective geometry. Their system only covers the penalty area and therefore they are not able to give a good insight of the entire game.

One of the most interesting systems, Aspogamo, is presented by Beetz, Michael et al. [3,4]. This system is able to analyze sports games using an ontology of models of the game that has as primitives the players' positions, motion trajectories and ball actions. The player detection is performed using first an intensity variance detection to isolate areas of interest and afterwards the players are identified using a combination of three colour templates: one for the shirt, another for the shorts and the last one for the socks of the players. Occlusions are identified using geometric constraints of shape and size. The multiple target tracking is done with a Rao-Blackwellized Resampling Particle Filter with fixed lag. This system is also able to track the ball using a particle filter algorithm.

Regarding indoor sports one of the most promising, mature and relevant work is Sagit developed by Pers et al. [5, 6]. They use a two fixed camera system placed in the ceiling of the sports hall that provides a bird's eye view.

A detailed explanation for camera temporal and spatial calibration and error analysis is given and the main objective of their work was to follow trajectories. Three different algorithms are compared and described, one based on motion detection, another on RGB (colour) tracking and a last one on a colour and template tracking. They apply their method to several sports such as handball, basketball and squash. However their system is much operator dependent and is used offline.

Several authors apply Kalman filter techniques to address some of these problems, for example Liu et al. [7] give a detailed explanation on their efforts to track skaters in a skating competition. They use a hierarchical model of two components: helmet – identified by a template matching approach and body – detected by a colour histogram matching method combined with an unscented Kalman filter.

Chris and Boyle [8] use a single camera system and apply a multiple object condensation scheme to a five player soccer game. Initially a sample (that corresponds to a bounding box) of each player is detected, then through a propagation algorithm the fitness of each bounding box is evaluated and adjusted. They also include an improved predictive stage that incorporates estimates of positions from a Kalman filter.

The significance of such systems to the area of knowledge of sports can be assessed by the work of Marko Sibila et al. [9] as they evaluate the importance of cyclic movements on handball. With the Sagit platform they were able to perform a study about identifying the differences of volume and intensity in large-scale cyclic movement activities performed by handball players.

In latter years there has also been a great development in robotics soccer and numerous papers were published [10–14] describing several features developed for image processing and tracking algorithms tested with success. Some of these features due to the similarity between indoor sports can also be explored.

3 Architecture

There are several types of cameras available in the market. Mainly for convenience reasons this work was based on a Sony SNCDM110 IP security camera. Recently, colour cameras based on GigEthernet interface have been introduced in the market therefore some considerations on a system based on these cameras are also drawn.

With this kind of technology it is possible to place the cameras far away from the processing system and it is quite easy to mount a multi-camera system using a standard, low cost Ethernet switch.

Due to the dynamics of indoor sports games, where players are constantly moving the best spot to place the camera is in the ceiling of the sports hall which gives a birds-eye perspective [5]. This way there will be no obstacles between the camera and each player and there will be a clear view of the ball, except when occlusion occurs due to a player.

The indoor sport that uses the biggest field area is handball. A handball field has 20×40 m which represents a very large area to be covered and also a very large amount of information to be processed and taken care of. Therefore special attention must be paid when choosing the cameras resolution and frame rates.

Using a system based on a moving camera would not be cheaper and would not prevent dark areas in the game that could eventually be of interest to a coach. The system would also be prone to aging and mechanical problems. Due to these limitations a fixed system seems to be the best choice.

3.1 Projected Solution

The vast area to be covered makes it impossible to use a single camera; therefore a four GigEthernet fixed camera system was projected. The camera placement as well as the system architecture is shown in Fig. 1.

The cameras chosen have resolutions of 640 × 480 pixels and maximum frame rates of 60fps. This implies a high quantity of data delivered by each camera and therefore the architecture is based on a two processor system.

3.2 Tested Solution

Due to some constraints it was not possible to implement the full system. Therefore, all the results presented in this paper were achieved using a single Sony SNCDM110 camera fixed above the 6 m handball line (Fig. 2).

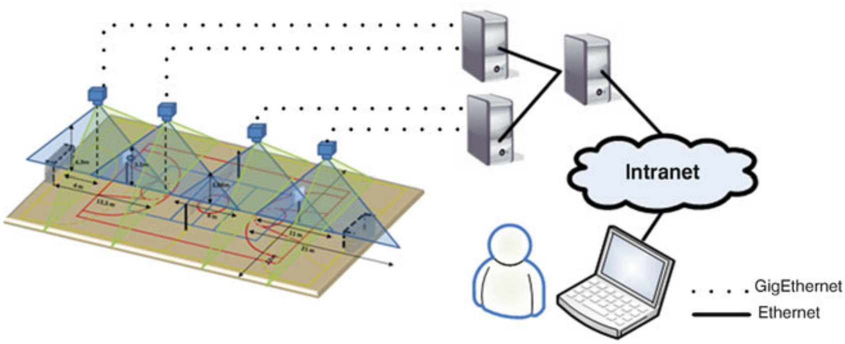

Fig. 1 System architecture

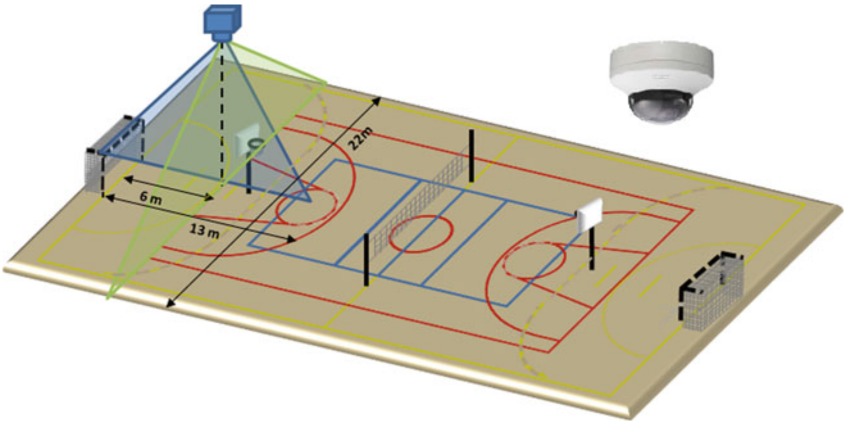

Fig. 2 Camera placement for the tested solution

4 Image Processing

One of the most important systems in this kind of application is the image processing. It must be robust enough so that the tracking algorithm can have solid bases to work with.

In this case the image processing system is based on blob definition and each player is seen as a colour blob. In order to have a clean interface with the images received from the camera we used the OpenCV library [15].

The following subsections describe the steps towards the players' detection.

4.1 Team Definition

The first step consists on defining the colour blobs for each team that correspond to sub-spaces of the entire colour space. This is a very important task since a good colour calibration will influence the success of the subsequent steps.

These colour sub-spaces are stored in a special lookup table that corresponds to a three dimensional vector with 32 elements in each dimension, this way each colour component (red, green and blue) is represented by the five most significant bits. The least significant bits of each colour component represent very slight changes in the colour itself, for that reason as well as for computational purpose they are ignored.

In this lookup table each point defined by the R, G and B components can have a few values, corresponding to a team colour or no team. The set of points belonging to the same team colour constitute a team colour sub-space.

Besides this, each team colour has a colour identifier that is used to mark a pixel that belongs to a given team colour.

Three approaches were tested [16], the first using the Red, Green and Blue (RGB) colour space and the other two using the Hue, Saturation and Luminance (HSL) colour space (which allow minimizing the shadow and light variations effects):

- *Physical range* – This is the simplest method tested and uses only the area selected by the user to define the team colour sub-space. The results showed that this is a very time consuming task that not always produces good results.
- *Physical flood* – Besides including in the team colour sub-space the area defined by the user (seed) it also includes the surrounding area as long as it has a colour similar to the one in the chosen area. The propagation among pixels will happen in all physical directions in a recursive way until reaching a pixel that has a colour too far away from the seed or from the previous neighbour. This novelty allows a faster colour calibration with very good results.
- *Colour grow* – Based on growing the colour space around the selected colour. With this method the colour expansion is not restricted by the physical neighbourhood and therefore gives a faster team colour blob definition.

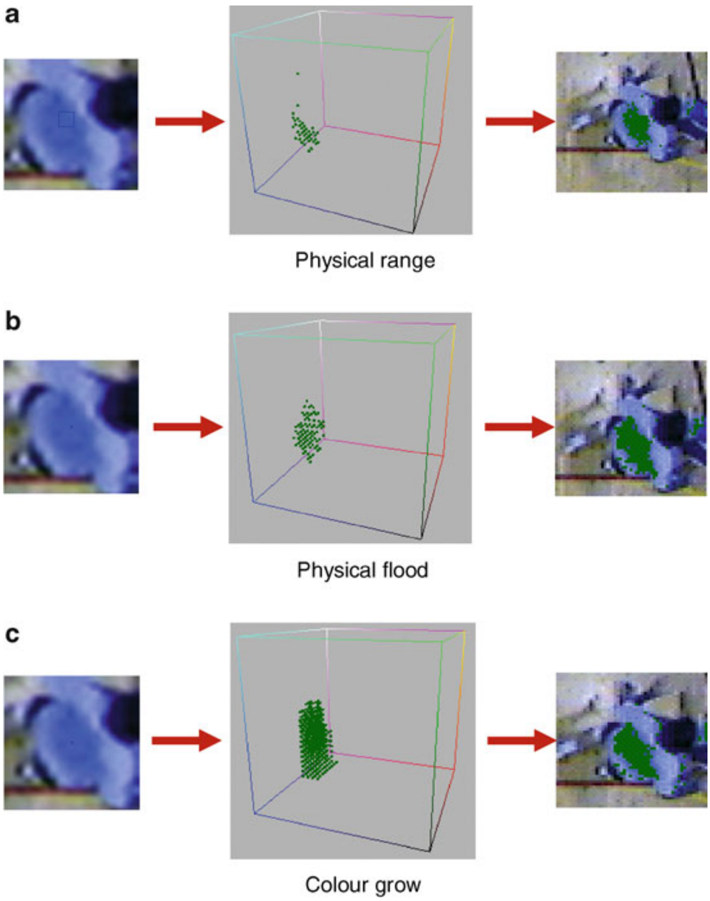

Fig. 3 Team colour sub-spaces resulting from different colour calibration techniques: (**a**) Physical Range, (**b**) Physical Flood, (**c**) Colour Grow

The colour team sub-spaces resulting from applying these three different techniques can be seen on Fig. 3. In this figure the team colour sub-space is identified with the green colour.

4.2 Background Subtraction

Most of the image area has non-useful information. In fact the regions of interest are the ones that include a player so it makes sense to perform a background subtraction in order to highlight the zones where the players are. To perform this subtraction an empty image of the field is collected a priori.

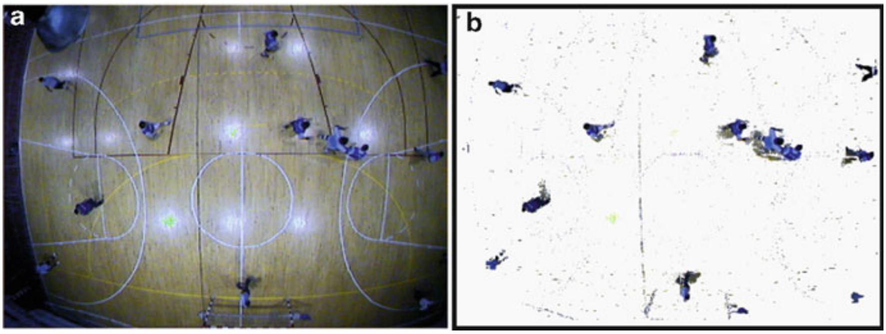

Fig. 4 Background subtraction before (**a**) and after (**b**)

Due to the brightness of the field empty image it is impossible to simple subtract it, because it will also have a huge impact on the players' figures almost eliminating them from the image.

Therefore a conditional subtraction per pixel is performed [16], which means that the subtraction occurs only if the pixel under analysis has a colour similar to the one of the background image. To be more specific the two pixels are not exactly subtracted but the pixel in the image under analysis will be updated with the white colour.

Figure 4 shows the effect of applying the background subtraction.

With this conditional background subtraction the areas containing the players are clearly highlighted.

4.3 Colour Detection

The first action towards the players' localization is to detect the colours of their uniforms. Since the colour sub-spaces for each team have already been calibrated and stored in the colour lookup up table it is necessary to scan the entire image to detect and mark pixels that belong to either one of those colour sub-spaces.

During this scan each pixel colour is tested to see if it has an entry on the colour lookup table. If it corresponds to a team colour then the colour of the pixel is replaced with the team colour identifier.

4.4 Blob Aggregation and Characterization

At this point there is only information if a pixel belongs to a given colour team or no team, it is still necessary to establish a relationship between pixels belonging to the same colour blob. A colour blob corresponds to a region, in the image, where pixels of the same colour team sub-space are concentrated.

The algorithm responsible for establishing this relationship comprises two steps. The first one is based on a per line scan detection and the information is stored in a way similar to that of a run-length encoding using three parameters: y, x_{min} and x_{max}. Whenever a pixel belonging to a team colour is reached the subsequent pixels of the same line are checked to see if they belong to the same team colour. An outline of the algorithm is presented next.

```
for y:=0 to imageMaxY
  for x:=0 to imageMaxX
    curColour:=getColourFromImage(x,y)
    if curColour<>-1 then
      Segment(SgmCount++).Colour:=curColour
      Segment(SgmCount).Start:=x
Segment(SgmCount).End:=findEndSegment(curColour,x,y)
      if Segment(SgmCount).End <>-1
        SgmCount++
      end if
    end if
  end for
end for
```

Where the findEndSegment function is described by:

```
findEndSegment(curColour,xStart,y)
  for x:=xStart to endOfLine
    while getColourFromImage(x,y)=curColour
      x++
    end while
    xEnd:=x
    while getColourFromImage(x,y)<>curColour
      x++
    end while
    if xEnd-x>BLOBDISTANCE
      break
    end if
  end for
  if xEnd-x>BLOBSIZE
    return xEnd
  else
    return -1
    end if
  end findEndSegment
```

As a result from the previous procedure a series of single lines are identified as belonging to a specific colour team and they still need to be joined to form a single blob, which represents the second step.

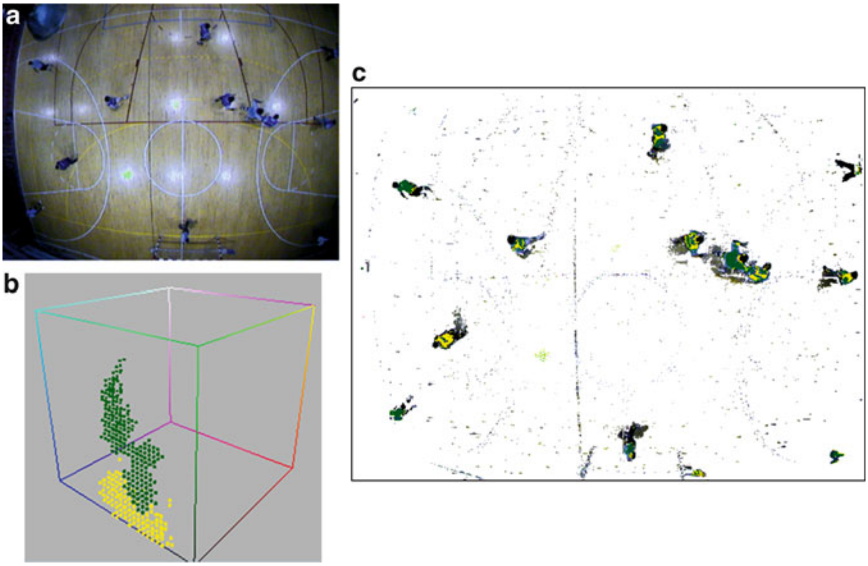

Fig. 5 Colour detection and blob aggregation. Original image (**a**), teams colour subspaces (**b**) and image after processing (**c**)

If the distance between two of these lines is small and they belong to the same colour team they are considered as being part of the same blob and are connected together.

Once the blob aggregation is performed it is possible to characterize each blob, namely determine its maximum and minimum x and y positions, the area it occupies, the rectangle that best fits the blob and its centre of mass.

Figure 5 shows the result of these processing techniques including the colour regions of each team.

Despite the two team colour sub-spaces being too near the players from each team can be properly identified as demonstrated by the bottom image of the above image.

Two features are used in order to determine if a colour blob can represent a player: the area of the colour blob and the colour density inside the rectangle that best fits it.

4.5 Real World Transformation

Once the players are completely identified and the blobs characterization is performed it is still necessary to perform the conversion into the real world coordinates.

This real world transformation comprises two parameters one, as seen from the previous images is the barrel distortion, that in this case is quite severe and the other corresponds to a scaling factor between the coordinates of the image (that are in pixels) into the coordinates of the real world (in centimetres).

In order to compensate for these two factors Eq. 1 was considered [16]:

$$\begin{bmatrix} x_R \\ y_R \end{bmatrix} = \left(1 + kr^2\right) \begin{bmatrix} S_x \times x_D \\ S_y \times y_D \end{bmatrix} \tag{1}$$

$(1 + kr^2)$ term denotes the barrel effect distortion, being k the barrel distortion coefficient. (S_x,S_y) represents the scale factor, (x_R,y_R) are the real world coordinates and (x_D,y_D) are the distorted coordinates that correspond to the image coordinates.

In this equation the parameters that need to be found are the k and the (S_x,S_y), so that the real world coordinates for any given pixel can be determined.

The (S_x,S_y) can be easily calculated aided by the lines of the field and using the ratio between the real world distance and the pixels distance. For example, measuring the pixels used to define the inner area of the 6 m line.

Using this method both on the x axis as well as on the y results in: S = (2.94 cm, 2.80 cm).

For determining the k factor an approach based on the least square method was used. Let's ignore for now the scale factor and using the y component of Eq. 1 an expression like the one in Eq. 2 would be obtained:

$$\begin{aligned} &y_D = -y_D r^2 k + y_R \Leftrightarrow y = mx + b, \\ &where \quad y = y_D,\ x = -y_D r^2,\ m = k \: and \: b = y_R \end{aligned} \tag{2}$$

Applying the least squares method results in:

$$\begin{aligned} m &= \frac{n \sum_{i=1}^{n} \left(-y_D r^2 \times y_D\right) - \left(\sum_{i=1}^{n} -y_D r^2\right) \times \left(\sum_{i=1}^{n} y_D\right)}{n \sum_{i=1}^{n} (-y_D r^2)^2 - \left(\sum_{i=1}^{n} -y_D r^2\right)^2} = k \\ b &= \frac{\left(\sum_{i=1}^{n} y_D\right) \times \left(\sum_{i=1}^{n} (-y_D r^2)^2\right) - \left(\sum_{i=1}^{n} -y_D r^2\right) \times \left(\sum_{i=1}^{n} -y_D r^2 \times y_D\right)}{n \sum_{i=1}^{n} (-y_D r^2)^2 - \left(\sum_{i=1}^{n} -y_D r^2\right)^2} = y_R \end{aligned} \tag{3}$$

To accomplish a good calculation of the k coefficient is good practice to choose a line that covers the entire field. Figure 6 depicts the outcome of applying the barrel "undistortion" equation.

After the coordinates are transformed into the real world it is possible to determine the exact amount a player has run and if during a move is in a good position.

4.6 Player Tracking

Once the players are detected it is possible to perform their tracking. As stated before the main parameters that characterize each player are the area and the centre of mass.

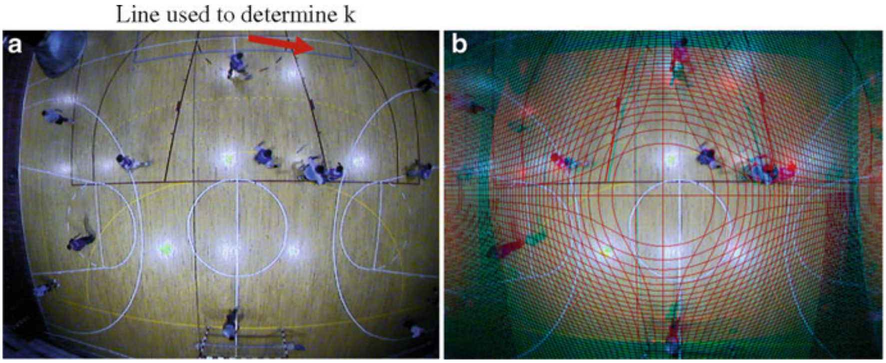

Fig. 6 Image before (**a**) and after applying the barrel "undistortion" expression (**b**)

The method adopted is based on the past information and on defining a probable area around each player that defines his/her next position. This probable area takes into account the position and the velocity of the player.

Once a new frame is picked, the processing system starts by identifying all the colour blobs and those that are identified as being a player are characterized.

The characterization parameters will be compared with the ones from the previously identified blobs and if they fit inside the probable area of one of those blobs then the new blob is assumed as being the sequence.

5 Results

This section presents the results achieved so far and gives a description of the test platform used.

Tests were conducted at the FADEUP's sports hall during handball training sessions of the Futebol Clube do Porto junior teams with training vests. Due to the short period the camera was available (the camera was lent by Sony Portugal) it was only possible to capture a few videos and not always under the best conditions.

5.1 Overview

Figure 7 demonstrates the final result after applying the image processing and the tracking techniques.

Merging between players of the same team seems to be the strongest factor to loose player identification.

Experiments also evidenced that the colour calibration of each team is a key factor to have a good player detection, and a combination of the physical flood and colour grow techniques produced the best results.

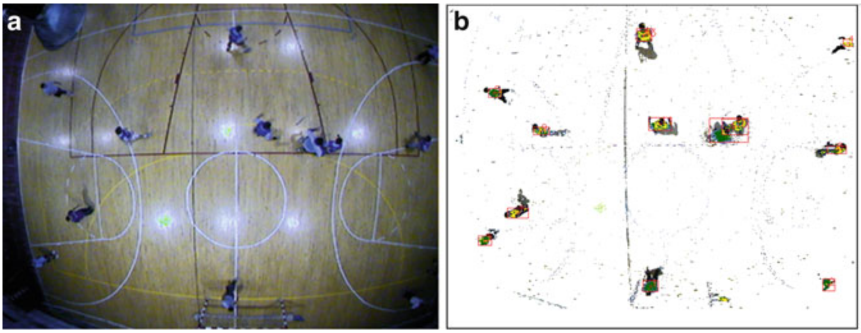

Fig. 7 Image before (**a**) and after (**b**) processing

Regarding the tracking algorithm the results are still scarce, mainly because we were only able to cover a small portion of the field and therefore when a player leaves this area the tracking algorithm is not able to follow him.

However when the player stays in the visible region the tracking behaviour is satisfactory and even when a player is not detected in one frame it is able to recover the track in the subsequent frames.

5.2 Sample Footage

The following results were obtained using a 60 s' portion (which corresponds to 900 frames) of a video recorded at a handball training session of the Futebol Clube do Porto junior team, filmed with permission at the Faculty of Sports of the University of Porto on the 25 April 2009.

The video was filmed in MJPEG with resolution of 640 $\times$ 480 pixels, 15 fps and compression ratio of 1/6.

The sample footage used has a particularity that is common on training sessions but not on real game situations, both teams have very similar equipments only differing in a dark blue vest the dark blue team wears above the same equipment of the light blue team.

5.3 Player Detection

Robust player detection plays a very important role since it establishes solid foundations where the tracking algorithm can build upon. Therefore the first tests were oriented to determine the player detection robustness.

The left side of Fig. 8 shows the team colour definition used in this batch of tests (green colour identifies the light blue team and yellow colour the dark blue team)

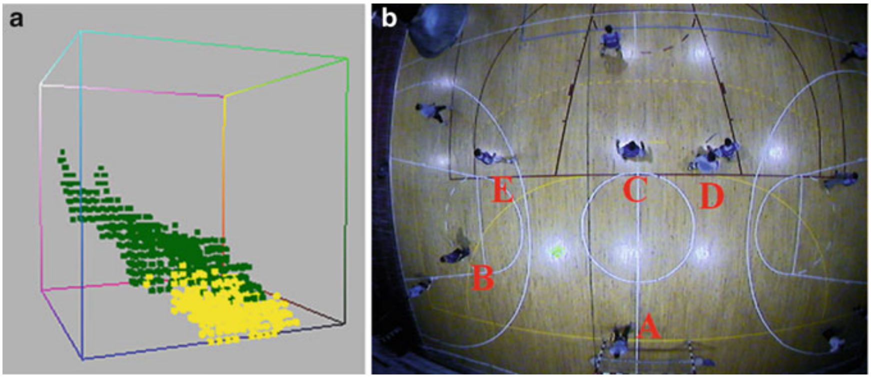

Fig. 8 Team colour definition (*left*) and players under detection (*right*)

Table 1 Player detection rate

Player	Detection rate (%)
A (goalkeeper)	76.2
B (dark blue team)	72.2
C (dark blue team)	88.0
D (light blue team)	93.3
E (dark blue team)	75.2

where it is possible to see that due to the equipment similarity the two team colours have an interlaced region which makes it even harder to identify players whose equipment colour is on this frontier.

On the right side of the same figure players used to evaluate the detection rate are highlighted and identified with letters

Table 1 provides the detection rates for each highlighted player.

From these results it is possible to see that players' detection rates ranged from 72.2% (player D) until 93.3% (player D). It is also noticed that the detection algorithm had better performance at the centre of the image because this area had a better illumination (there were some missing light bulbs on the sides of the field mainly near the goal) and also corresponded to the area seen by the centre of the lenses where the image is better.

Player B had the lowest detection rate (72.2%) followed by player E (75.2%). The first case is explained by the fact that, during the sample footage, the player stayed most of the time in the extremity of the image where the two above mentioned conditions were not gathered. On the second case the player was from the dark blue team but stayed the majority of time with his flank towards the camera where the most visible area of the equipment was the sleeve (of light blue colour) and therefore the player was miss classified as belonging to the wrong team.

Player C was also from the dark blue team but had a satisfactory detection rate because he consistently stayed on the centre of the image with his backs towards the camera.

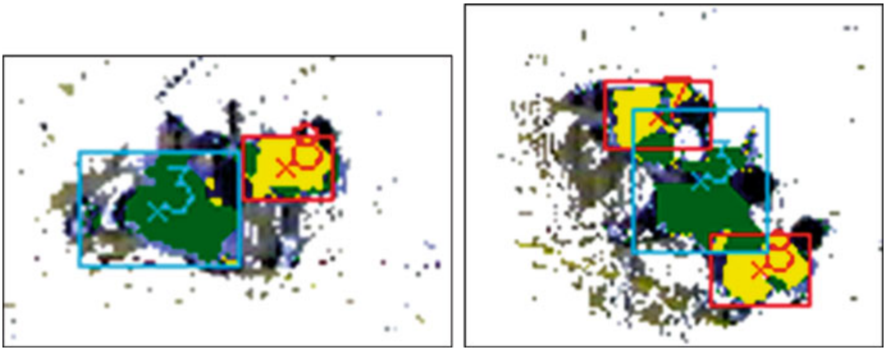

Fig. 9 Correct player identification during a merge situation

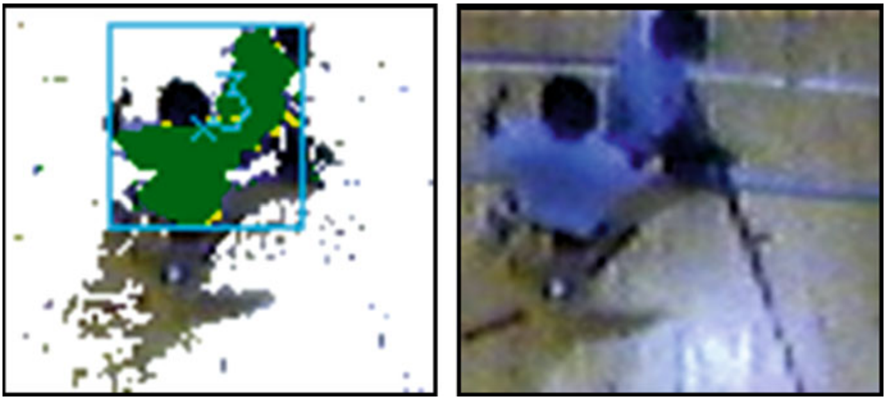

Fig. 10 Merging between players of the same team

It was also possible to verify that a player belonging to the light blue team (player D) had better chances of being well detected

Further tests showed that during merge situations with players from different teams the detection was able to perform correctly as can be seen on Fig. 9.

Nevertheless when two players of the same team got close enough to occur merging (Fig. 10) the detection algorithm considered them as being a single individual.

The image on the right side shows clearly that the camera "saw" the sleeves from both players touching.

5.4 Player Tracking

The performance of the tracking algorithm was evaluated by the time it continuously managed to track a player.

As stated before a good tracking algorithm is based on solid player detection, therefore a lower detection rate will imply a worse player tracking which is proved by the results on the following figures and tables.

Tables provide information about how many times the tracking algorithm lost a player identified by a higher or lower number of sections. Each section represents a period of time in which the tracker was able to continuously follow the player. Information about the distance run by the player is also given.

Figures provide a graphical visualization of the area covered by the player in each section (each colour identifies a section tracked properly). Although in reality the set of sections belong to the same player the tracking algorithm sees each portion as belonging to a different player.

We present the results achieved for player B and C (recall 5.3).

Figure 11 and Table 2 represent the data for player B. Initially the player was tracked in section dark grey, but 14.2 s later the player was not detected for a long enough period and his track was lost. During this sample footage the tracking algorithm lost the player three times.

We could achieve better results on player C as shown on Fig. 12 and Table 3. The tracking algorithm only lost the player two times and on the second section it was able to follow the player during 30 s.

As stated before, merging and occlusion can carry many troubles in a tracking algorithm, nevertheless tests showed a good performance when this happens between players from opposite teams.

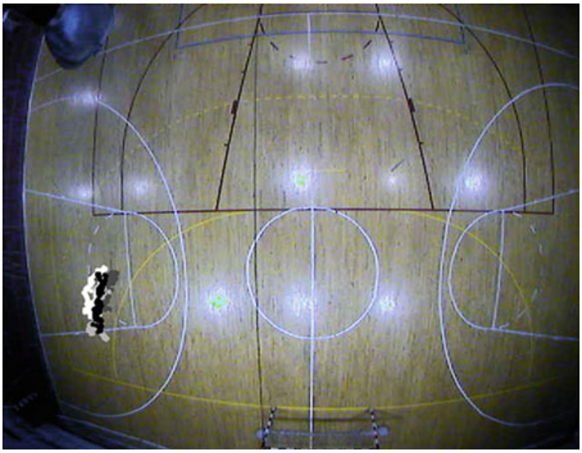

Fig. 11 Spatial occupation of player B

Table 2 Player B sections during tracking

Section	Duration (s)	Distance (cm)
Dark grey	14.2	12.3
Light grey	14.4	8.1
White	16.2	9.7
Black	12.2	8.3

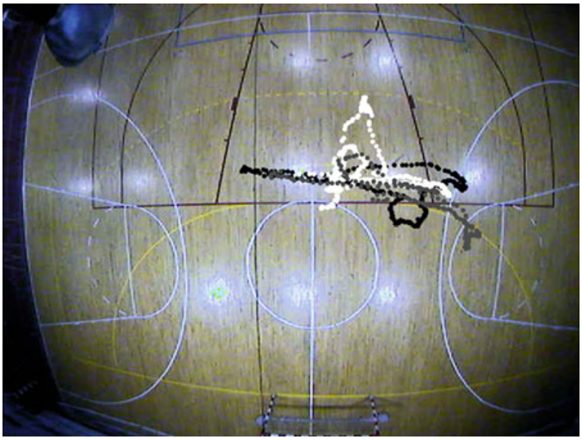

Fig. 12 Spatial occupation of player C

Table 3 Player C sections during tracking

Section	Duration (s)	Distance (cm)
Black	16.6	23.2
White	30.0	31.3
Grey	13.1	21.8

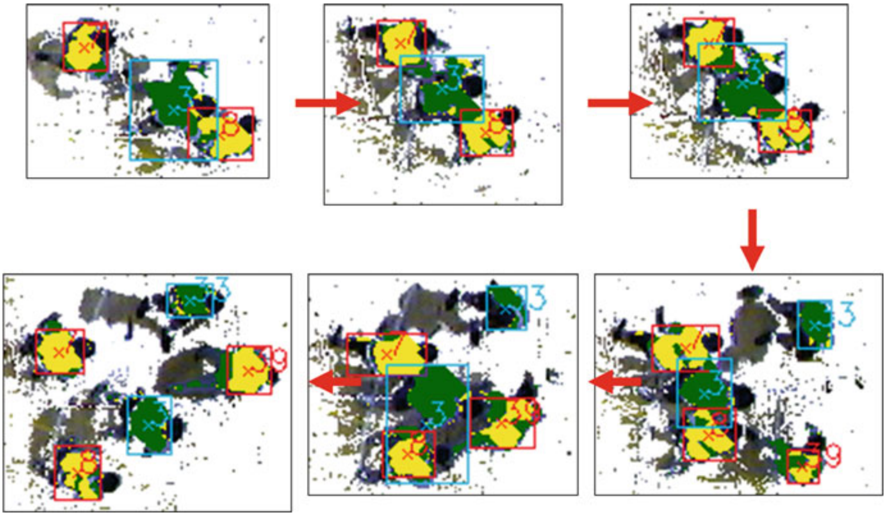

Fig. 13 Tracking during a merging situation

Figure 13 shows the result of the tracking algorithm during the right side merging of Fig. 9. This sequence of screenshots evidences that, despite occurring a merging between several players, the algorithm was capable of identifying correctly the players during and after the merging period.

Of course that during the merging some players were seen much bigger (example of player three), but in the general sense it is able to detect all the players and position their centre of mass (cross inside the rectangle) in more or less the correct position.

These tests were performed in a laptop computer with 1MB L2 cache and powered by an Intel T2130 processor running at 1.86 GHz, under Windows Vista operative system and, in average, each frames takes around 65 ms to be processed.

6 Conclusions and Future Work

This chapter presented a system for tracking players in indoor sports games. The main objectives were to develop a system that could handle such a complex problem, recognize each team player and follow their movement all over the field.

The system is composed of two main blocks, one concerned with image processing where several colour techniques are applied in order to identify the players and the other responsible for tracking the players throughout the field.

Tests conducted using an IP surveillance camera and a sample footage of 60 s showed that most of the time players are correctly identified has being part of either team. This is a good result taking into consideration the several artefacts the camera produces and the many light effects present on this scenario (mirror of the lights, shadows of the players, and barrel effect among others).

The results also evidence that players detection based on blob notion achieves good results, although it is not completely robust to player's merging and occlusions. The best result achieved was 93.3% of correct identification and the worst 72.2%.

Finally the last step, player tracking, was based on defining a probable area around the previous position of the player using the centre of mass position and the maximum theoretical player velocity. The minimalistic presented tracking algorithm was able to correctly and continuously follow a player during 30 s.

Despite the tests being conducted in a handball environment the application is generic and therefore should be able to analyse other sports images such as basketball or volleyball.

It is clear that there are many opportunities to improve and expand the work developed so far. Maybe one of the most important aspects to improve is the tracking algorithm in order to reduce its dependency on a good player detection. Therefore Kalman filter techniques and artificial intelligence methodologies may be explored in order to improve the algorithm's performance.

It would also be interesting to implement the complete engineering solution and test it in a real game situation where the players have very distinct equipments (with official team colours).

Acknowledgments The authors would like to thank Sony Portugal for the cost free usage of the camera used for the tests. The presented work was also partially funded by Fundação Gulbenkian through a PhD scholarship and by Portuguese FCT, under project PTDC/EIA/70695/2006 – ACORD – Adaptative Coordination of Robotic Teams.

References

1. Liu, J., Tong, X., Li, W., Wang, T., Zhang, Y., Wang, H., Yang, B., Sun, L., Yang, S.: Automatic player detection, labeling and tracking in broadcast soccer video. Pattern Recognit. Lett. **30**(2), 103–113 (2009)
2. Iwase, S., Saito, H.: Tracking soccer players based on homography among multiple views. In: Proceedings of SPIE – The International Society for Optical Engineering, Lugano, Italy **5150**(3), 283–292 (2003)
3. Beetz, M., Gedikli, S., Bandouch, J., Kirchlechner, B., von Hoyningen-Huene, N., Perzylo, A.: Visually tracking football games based on TV broadcasts. In: Proceedings of the 20th International Joint Conference on Artifical intelligence, pp. 2066–2071. Morgan Kaufmann Publishers Inc., San Francisco, CA (2007)
4. Beetz, M., von Hoyningen-Huene, N., Kirchlechner, B., Gedikli, S., Siles, F., Durus, M., Lames, M.: ASPOGAMO: Automated sports games analysis models. Int. J. Comput. Sci. Sport **8**(1), 4–12 (2009)
5. Pers, J., Kovacic, S.: Computer vision system for tracking players in sports games. In: Proceedings of the First International Workshop on Image and Signal Processing and Analysis, 2000, pp. 81–86. Pula, Croatia (2000)
6. Pers, J., Bon, M., Kovacic, S., Sibila, M., Dezman, B.: Observation and analysis of large-scale human motion. Hum. Mov. Sci. **21**(2), 295–311 (2002)
7. Lui, G., Tang, X., Huang, J., Liu, F., Sun, D.: Hierarchical model-based human motion tracking via unscented Kalman Filter. In: Proceedings of the 11th International Conference on Computer Vision, pp. 1–8. IEEE Computer Society Press, Washington, DC (2007)
8. Needham, C.J., Boyle, R.D.: Tracking multiple sports players through occlusion, congestion and scale. In: Proceedings of the British Machine Vision Conference 2001, vol. 1, pp. 93–102. Manchester, UK (2001)
9. Šibila, M., Vuleta, D., Pori, P.: Position related differences in volume and intensity of large scale cyclic movements of male players in handball. Kinesiology **36**(1), 58–68 (2004)
10. Sousa, A: Arquitecturas de Sistemas Robóticos e Localização em Tempo Real Através de Visão. Feup Edições, Porto, Portugal (2003)
11. Moreira, A., Sousa, A., Costa, P.: Vision based real-time localization of multiple mobile robots. In: 3rd International Conference on Field and Service Robotics, pp. 103–106. Helsinki, Finland (2001)
12. D'Andrea, R., Lee, J.W., Hoffman, A., Yahaya, A.S., Cremean, L.B., Karpati, T.: Big Red: The Cornell small league robot. Lecture Notes in Computer Science 1856: 89–98 (2000)
13. Reis, L.P., et al.: Robust vision algorithms for quadruped soccer robots. In: Proceedings of CompImage 2006 – Computational Modelling of Objects Represented in Images: Fundamentals Methods and Applications, pp. 367–372. Coimbra, Portugal (2007)
14. Reis, L.P., Neves, A., Sousa, A.: Real-Time Vision in the RoboCup – Robotic Soccer International Competitions, Computer Vision in Robotics – VipIMAGE 2009 – II ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing, pp. 319–324. Porto, Portugal (2009)
15. Bradski, G., Kaehler, A.: Learning OpenCV. O'Reilly Media, USA (2008)
16. Sousa, A., Santiago, C., Reis, L.P., Estriga, M.L.: Automatic detection and tracking of handball players. In: Proceedings of the VipIMAGE 2009 – II ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing, pp. 213–219. Porto, Portugal (2009)

http://www.springer.com/978-94-007-0010-9

Computational Vision and Medical Image Processing
Recent Trends
Tavares, J.; Jorge, R.N. (Eds.)
2011, IX, 349 p., Hardcover
ISBN: 978-94-007-0010-9