AdaBoost in basketball player identification

Z. Ivankovic*, B. Markoski*, M. Ivkovic*, D. Radosav*and P. Pecev**
* University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia
** University of Novi Sad, Faculty of Science, Novi Sad, Serbia
zdravko.ivankovic@tfzr.rs

Abstract — **Video materials contain huge amount of information. Their storage in databases and analysis by various algorithms represents an area that constantly develops. This paper presents the process of analysis of basketball games by AdaBoost algorithm. This algorithm is mainly used for face recognition and body parts recognition. It consists of a linear combination of weak classifiers. In this paper, stumps were used as weak classifiers. The aim of this research is to assess the accuracy of this algorithm when applied in players' identification at basketball games. Capabilities of AdaBoost were examined when applied to video footage from single moving camera, and when these footages were not previously treated by any other algorithm. The first training was performed by entire images of basketball players, whereas the second training was performed by using the images of the head and torso. By applying the algorithm to the given set of images that include head and torso, the algorithm obtained an accuracy of 70.5%. Experimental results have shown that training on the set of entire body images was not possible due to large amount of background that goes into the training, and which represents noise in training process. This accuracy could be increased by applying filters that would remove background from images and leave just basketball players. By applying those filters, the amount of noise in the training data would be significantly reduced.**

I. INTRODUCTION

With the advance of information technology, the amount of multimedia content that is created, transmitted and stored constantly increases. As a result, the multimedia content is widely used in many applications. Therefore, there is the necessity for its organization and analysis, both from commercial and academic aspects. Computer vision represents a technology that can be applied in order to achieve effective search and analysis of video content.

Computer vision represents technology with constant growth and progress in recent years. Its application includes a number of very important life areas such as medicine (examination of medical records and their automated analysis), public safety (security camera examination and person recognition), traffic safety (detection of pedestrians, cyclists, animals that are located next to road, driving line and driver sleepiness), analysis of sports event (recognition of players, their movements and tactics they apply), etc.

Computer vision represents a process that consists of several phases [1]. The first phase is initialization, which represents the process of removing the background and extracting objects of interest by creating their models using markers, images or predefined shapes. The next phase is tracking, which represents the process of object recognition in successive frames. This phase lasts until the object leaves observed area, or until the tracking has been terminated. The third phase is pose estimation, which in the process of human recognition represents analysis of the arms, legs, body and head, according to which the object is classified into one of the previously defined poses. The final phase is recognition that can be achieved by recognizing person's face or some other characteristic feature.

This paper presents application of computer vision in identifying players in basketball games. Players are identified in the videos that are distributed to a wide range of viewers via television stations. These videos show basketball court and players by only one camera in particular point of observation. Negative aspects of the analysis performed using videos for TV broadcasting represents the fact that certain actions appear more than once (repeated shots), and that certain shots contain things that have no direct connection with the observed match (shots of audience, event announcements, interviews with celebrities or players). In addition to this type of analysis, in practice there are two more methodologies. The first one is basketball players identifying by using markers placed on them, which is the most accurate way of identifying and tracking. The second way is analysis by using multiple synchronized cameras placed on pre-planned positions on the basketball court. At least one of these cameras makes every player fully visible at any time. In the analysis of basketball games broadcasted by TV stations, there is often a situation that one player is obscured by other players. In this situation, we must make an approximation of his position according to the pose of other players in his vicinity. This inevitably introduces a degree of imprecision in the analysis itself. The second and third ways of analysis are more precise methodologies that require special conditions and special hardware.

Computer vision, therefore, consists of four phases which are described shortly before (initialization, tracking, pose estimation and recognition), and the subject of this study is the first phase, e.g. initialization, which represents creating of the model of basketball players based on their images. Images are obtained using specially developed tools, which store one frame from video material in every 0.5 second on provided location on computer.

Rectangles that contain basketball players are cut from these images and they will be used in training process. Training is done with AdaBoost algorithm, which represents an often-used algorithm in the shape recognition. It is primarily used for face recognition and body parts recognition. The aim of this study is to assess the capabilities of this algorithm to identify players at basketball games. The images of basketball players are objects with high degree of diversity, depending on whether the player has the ball, plays defense, shoots for

978-1-4673-5206-2/12/$31.00 ©2012 IEEE

the basket, jumps for the ball etc. Therefore, the training set has many differences, and the goal of this research is to find the answer to the question whether AdaBoost can be successfully used in such training sets. Its main feature is the speed, which is especially important when analyzing basketball games, because coaches often want to have complete analysis of the game during the game, in order to make some changes in its team play. Therefore, the use of this algorithm in the analysis of games would be useful. AdaBoost requires a large training set. Therefore, six thousand positive examples (images of basketball players) and six thousand negative examples (images that do not contain any basketball player) were used in each training process. These images are combined in the training process in such a way that images of basketball players (that can be resized to some extent, or certain curve degree can be applied) are added to precisely defined locations in the pictures that do not contain basketball players. Images combined in this way are used in training process to enable assessment of the quality of AdaBoost algorithm.

II. Related Work

Data mining in sport is experiencing rapid growth in recent years and is slowly attracting the attention of the largest sports associations. The baseball team Boston Red Sox and the football club AC Milan were among the first organizations that started to apply the benefits of data mining. People with special merits for the introduction of data mining in the sport were Dean Oliver, who introduced this methodology in basketball and Bill James, who did the same in baseball.

According to Schumaker et al. [2] in the next few years, the application of data mining in sports will face several challenges and obstacles. The biggest obstacle will be to overcome opposition to new technologies that some members of sports organizations, who prefer traditional way of acquiring knowledge. The same authors suggest that the application of data mining in sport is at a critical point and that a number of features it brings are waiting to be exploited. Some of these features will quickly bring the desired results, while others will take years and even decades. They also point out that the primary task is not to determine the correct way in collecting data, but to determine which data should be collected and how to use them in the best manner. Markoski et al represented some basic ways of using data collected in basketball games [3].

While the use of statistics in decision-making is certainly an improvement over the use of instinct of coaches, managers and scouts, statistics alone can go in the wrong direction without knowledge of the problem domain. The first part of the problem is to determine the performance metrics. A large number of existing sport metrics can easily be used inappropriately [4]. Ballard [5] has presented a typical example of inaccuracy in data collection in basketball. He gives the example of a jump in the defense, which represents the number of times a player catches the ball in the defense after the opponents unsuccessful shot. In order to record a jump in the defense, teammates have to block the opponent's players and keep them away from the basket. By blocking the opposing players, basketball players usually do not have the opportunity to catch the ball. However, their performance in the defense makes them equally important in order to catch the ball. By observing the way of recording defense rebounds, it can be seen that only the player who catches the ball is awarded with rebound. The second part of the problem is to find interesting patterns in data. These patterns may display movement and intentions of opposing players, reveal the beginning of injury during training or predict outcome based on the previous games. Practical method in finding those patterns could be appliance of neural networks [6].

Analysis of human activities using computer vision is widespread. Its attractiveness is based on a broad area of application and great complexity. A large number of systems for monitoring, based on computer vision, imply the fact that manlike kinematic structure with a fixed number of joints and known degree of freedom of movement already exists. Initialization is then limited to the assessment of limb length. Commercial systems use markers that record pre-defined movements. They are used to determine the individual degrees of freedom of movement. Known correspondence between the markers and limbs, together with the reconstructed 3D markers trajectory during the movement is used to accurately determine the length of the limbs. Strong restrictions on the symmetry of left and right sides of the skeleton are often imposed in the calculations. A large number of papers are related to body pose initialization and limb length [7][8][9].

Direct estimation of kinematic structure based on several moving people is also the subject of research. Krahnstoever et al. presented a method for automatic initialization of kinematic structure of the upper body, based on a movement extracted from a series of individual pictures [10]. Songet et al. presented unsupervised learning algorithm, which follows the characteristic points in complex scenes in order to create a kinematic model of the entire body automatically. Learned models are then used to track the movement of pedestrians by laterally placed cameras [11]. These approaches provide general solutions to the problem of kinematic model initialization because they directly distinguish the desired structure from the observed scene.

There are methods that are used to extract the kinematic structure from a variety of 3D shapes. Cheung et al. initialized the kinematic structure person based on visual frame, by moving each joint independently. The whole skeleton is obtained by fitting the parts of the body in motion with the visual frame in a particular position [12]. Manier et al. presented an automated approach to 3D pose estimation based on the central axis of the person's visual frame. Kinematic structure is initialized independently in each frame, allowing robust tracking [13]. Brostow et al., [14] presented a more general framework to assess the structure of the spine based on 3D shape time sequence. The spine is estimated in each frame based on the known forms that have been identified for this purpose. These papers show the approximate reconstruction of kinematic structures for babies, adults and animals.

Initializations of angles in the joints and limits they define in kinematic structure of the human body represent an important constraint that must be considered in order to estimate possible movement and poses. Manual specifying of the angles of the joints is a common method used in a number of algorithms for motion estimation using anthropometric data. Complex nature of the restrictions in the joints and coupling between restrictions for different degrees on freedom of movement is not taken into account. Numerous researches which examine learning models for the restrictions in the joints and their correlation have been conducted in order to overcome these limitations. Anthropometric models for relationships between angles in the joints of the hands (shoulders, elbows and wrists) were used to obtain limits in visual tracking and 3D pose estimation of the upper body [15]. There have been conducted some studies that examine modeling of the joint limits based on measurements of human movement recorded by markers [16] and clinical data [17].These papers have demonstrated improved performance in the human pose assessment of a person that performs complex movements of the upper body.

Later studies have been directed to algorithms training in order to identify possible locations for parts of the body which are then combined probabilistically to locate people [18][19]. Initialization of such models requires a large set of training data for both positive (containing the object) and negative (do not contain the requested object) for different parts of the body. AdaBoost-like approaches have been successfully used to learn the parts of the body such as face [20], hands, arms, legs and torso [18].

The sequences of human motion tracking from commercial systems that use markers are increasingly used in order to learn models of human kinematics and specific movements, all in order to provide constraints for subsequent researches in this area. Databases with recorded movements have also been used in order to create a synthesis of images with known 3D poses. The goal is to identify knowledge needed to map images into the poses.

III. AdaBoost Algorithm

Boosting takes its origin from the theoretical framework for studying machine learning called "PAC" (Probably Approximately Correct). The question was whether "weak" learning algorithm, which behaves slightly better than random guessing, could be a building block for general accurate "strong" learning algorithm.

AdaBoost, short for Adaptive Boosting, is a machine-learning algorithm, which was first formulated by Freund and Schapiro [21] solved many practical difficulties that previous boosting algorithms have encountered with. AdaBoost is adaptive in the sense that classifiers that come in the execution are adjusted in the direction of those instances that were wrongly classified with the previous classifiers. It is sensitive to the noisy data and information that does not belong to the required set. However, in some situations, this algorithm may be less susceptible to memory input set in comparison to many other algorithms. AdaBoost calls the weak classifiers repeatedly performing a series $of t = 1, \dots, T$ out of T classifiers. For each call of weighted distributions, D_t is updated to point to the importance of the examples contained in the data set for classification. In each series of execution, weight of incorrectly classified examples increases (or, alternatively, weights of each correctly classified example are reduced), so that the new classifiers focus more on those examples. This is a meta-algorithm and it can be used together with a number of other algorithms to improve their performance. Pseudo-code for AdaBoost is given below.

For a given: $(x_1, y_1), \dots, (x_m, y_m)$ where $x_i \in X, y_i \in Y = \{-1, +1\}$, initialize $D_1(i) = 1/m$.

For $t = 1, \dots, T$:

- Train weak learner using distribution D_t
- Get weak hypothesis $h_t: X \to \{-1, +1\}$ with error
$$\epsilon_t = Pr_{i \sim D_t}[h_t(x_i) \neq y_i]$$
- Choose:
$$\alpha_t = \frac{1}{2} \ln\left(\frac{1 - \varepsilon_t}{\varepsilon_t}\right)$$
- Update:
$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t} \; ako \; h_t(x_i) = y_i \\ e^{\alpha_t} \; \; ako \; h_t(x_i) \neq y_i \end{cases}$$
$$= \frac{D_t(i) exp\left(-\alpha_t y_i h_t(x_i)\right)}{Z_t}$$

where$Z_t$is a normalization factor (choset so thatD_{t+1} will be a distribution). Output the final hypothesis:

$$H(x) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)$$

The algorithm as input receives a training set $(x_1, y_1), \dots, (x_m, y_m)$ where each x_i belongs to a particular domain or instance space X, and each label y_i is in a label set Y. In most cases, it is assumed that $Y = \{-1, +1\}$, except when looking at extending of AdaBoost-with more classes. AdaBoost calls “weak” learning algorithm repeatedly in execution steps $t = 1, \dots, T$. One of the basic ideas of the algorithm is to maintain a distribution or set of weights over the training set. Weighting distribution in training example i in step t is denoted by $D_t(i)$. At the beginning, all the weights are placed on the same value, but at each step, the weights of incorrectly classified examples are increased so that the weak learning algorithm is forced to focus on more difficult examples in the training set.

The task of the weak learning algorithm is to find a weak hypothesis $h_t: X \to \{-1, +1\}$, which corresponds to the distribution D_t. The accuracy of the weak hypothesis is measured by its error:

$$\epsilon_t = Pr_{i \sim D_t}[h_t(x_i) \neq y_i] = \sum_{i: h_t(x_i) \neq y_i} D_t(i)$$

The previous expression shows that the error is measured in relation to the distribution D_t over which the weak learning algorithm was trained. In practice, the weak learning algorithm can be algorithm that uses weights D_t from the training set.

Looking at the example of recognizing players in basketball games, x_i is a representation of a player (standing motionless on the court, shooting to the basket, jumping for the ball, playing defense, attempting dribbling penetration, ...), while labels y_i show whether a given representations are basketball players or something else in the video. A presumption that certain objects are the players represents weak hypothesis, and sub collections that hypothesis examine are selected according to the distribution D_t.

When it comes to the hypotheses h_t, AdaBoost determines parameter α_t. Intuitively, α_t measures importance assigned to the hypothesis h_t. Previous listing shows that $\alpha_t \geq 0$ if $\epsilon_t \leq 1 / 2$, and that α_t increases its value as ϵ_t becomes smaller.

Next step is to update distributions D_t by using rules shown in previous listing. The effect of this rule is to increase the weight of examples that are misclassified by the hypothesis h_t, and to reduce weights of well-classified examples. Thus, the weights are trying to concentrate on "harder" examples.

The final hypothesis H is weighted majority of votes of T weak hypotheses, where α_t represents weight assigned to hypothesis h_t.

IV. Creating Model of Basketball Players

In order to create model of basketball players, AdaBoost algorithm was applied directly over the training set, without any preprocessing. During the training process of a new classifier, it is necessary to go through several stages:

1. Image acquisition
2. Example creation
3. Training

A. Image acquisition

Different sources can be used in the process of image acquisition. Since the aim of this paper is to identify players at basketball games, training process is using videos broadcasted by television stations. These videos are stored in multimedia databases. In order to apply AdaBoost algorithm additional software was used. The purpose of this software is to capture one frame from the video material in every 0.5 seconds and store it in predefined location on user's hard disk. That software is SampleCreator, which is implemented, in C# programming language and .NET 4.0 framework. Frames are captured and stored using DirectShow technology. If we assume that the game lasts 48 minutes (four quarters of 12 minutes in the NBA) and in addition to playtime, there is a number of interruptions (fouls, time outs, breaks between periods) we can assume the empirical value that the average duration of the match is about 100 minutes. By applying this method of storing frames from one game, we get 100 * 60 * 2 = 12000 frames. This provides a sufficient amount of data to train AdaBoost algorithm. This algorithm works with black and white pictures so color of shirt itself is not a major problem and an already trained algorithm can be applied in other games.

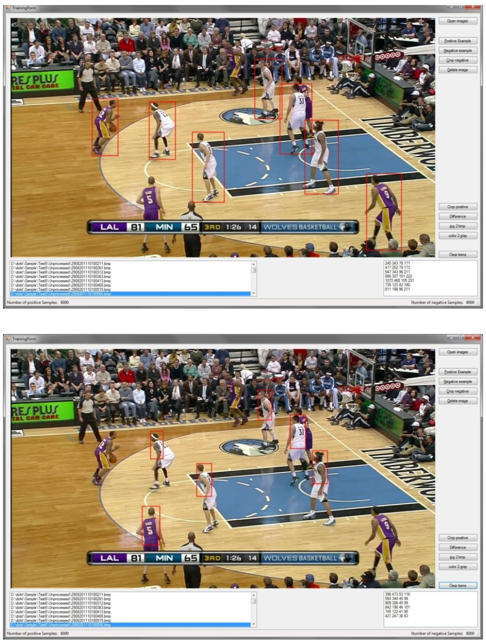

Figure 1. Creation of training examples

By using DirectShow technology, we obtain images that will mainly serve as positive examples (images that contain objects of interest). However, during the game occur frames that do not contain any player. These frames can be used as negative examples. In addition, negative examples can be found in other sources. It can be any pictures without basketball players.

B. Example creation

Examples are objects of interest, which are used in the training process in order to be found on the observed images. Depending on the required objects, different authors have used different sizes of training sets: 5000 examples in face detection [22], 6000 examples in pedestrian detection [23]. Training set creation is performed by cutting objects from the starting image. SampleCreator software is used for this purpose. It allows marking all the objects, while maintaining the ratio between width and height. Fig. 1 shows the process of marking the entire basketball players that will be used in training, and also the process of marking player's upper body. Different ratios of height and width were used in marking process. In marking the whole players, empirically was determined ratio of 2.2 (e.g. with 100px, height 220px), while in marking the upper body, that ratio was 1.8 (e.g. with 100px, height 180px).

Fig. 1 shows that not all players are marked as positive examples. The reason is that some objects may adversely affect the training if they are not shown clearly, if they are entering or leaving the frame, or if other objects obscure them. If we compare two images in fig. 1, it is evident that we have not marked the same players. Some players are not marked because of the background that would enter in the training process, which would likely adversely effect on AdaBoost algorithm. Unmarked players will not interfere in the training process because we crop positive examples from the picture and "glue" them on negative examples at pre-defined locations.

C. Training

AdaBoost algorithm is used in the training in order to be applied over the previously marked examples. Regarding training set and size of examples in that set, Kuranov et al. [22] have shown that the best results are achieved when the dimensions of examples that contain faces are reduced to 20x20 pixels. In the training process with images that contain whole basketball players, the ratio between the width and height cannot be 1:1. Therefore images were reduced to 20x44 pixels in training with the entire basketball players (574,479 features), and 20x36 pixels for training with players upper body (392,394 features).

The same authors have also suggested 20 stages training process. If as a training parameters we use degree of false positive of 0.5 and detection rate of 0.999, after the entire training we can expect degree of false positive of $0.5^{20} \approx 9.6\text{e-}07$, and the detection rate of $0.999^{20} \approx 0.98$.

During AdaBoost training on the examples that have symmetry (human face), we can apply some type of optimization that significantly speeds up processing. The reason is in the fact that in these cases only one-half (left of right) of the Haar feature is used. However, although players are objects that have symmetry, when observed during the game and in all positions in which they could be found, the application of symmetry in the training would not lead to desired results. In order to achieve higher accuracy, we used the extended set of Haar features. The basic set uses only upright features, while in this paper the full set of vertical features and features rotated by 45 degrees are used.

GAB (Gentle AdaBoost) classifier is used in the training. The work of Kuranov et al. [22] in which they have proven that GAB algorithm achieves highest results in object detection represents the main reason for application of this classifier This classifier is also the fastest one when considered time required for training.

During the training process, it has been used six thousand positive examples and six thousand negative examples. Training can be completed in some of sub stages when the minimal desired degree of search is fulfilled, or when the degree of false positive is reached, because the additional phases will certainly reduce this level (0.99 after current phase * 0.99 for next phase = 0.981 after next phase). Another reason for the completion of training is that all examples are rejected as incorrect.

During the training process over the set that included whole basketball players, the algorithm failed to reduce the level of false positive below 0.5.The algorithm continues training until the percentage of the original examples used in training falls to 0%, i.e. until all the examples are used. This would lead to interruption of the training without the wanted result. The reason that is behind this is, primarily, to a large degree of diversity that is encountered in the training set. This diversity is reflected in the position of players depending on whether they are standing, walking or running, and whether they have the ball, play an active defense or shooting on the basket. Another reason is a relatively large amount of background, which is located in the training examples, and which represents a noise.

Training on the set that contains players' upper body was successfully finished. This set does not include players' arms and legs and is therefore far more balanced. In addition, the amount of background in these images is far lower, as well as the noise.

V. Experimental Results

In order to assess the performances of a trained classifier, a set of testing images was used. Those images have precise locations for each object of interest. Tool that is used to assess performance, as input receive a collection of tagged images over which the classifier is applied. The performance of output is obtained by the number of found objects, the number of objects that are not found and the number of objects that are incorrectly classified as positive.

In this paper performances of the training over the set that contains upper body of basketball players are measured, because the training over the set that contains whole players was not possible. With the obtained results, a ROC (Receiver Operating Characteristic) curve was

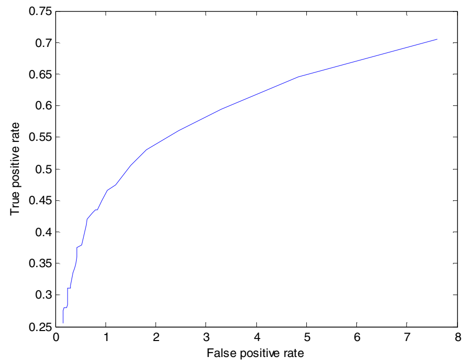

Figure 2. Detection performance

created. It graphically represents sensitivity, i.e. ratio of true positive versus false positive, for a binary classification system.

The output of the testing tool shows the number of objects that are identified (Hits), the number of objects that are not identified and which represents false negative (Missed), and the number of objects that are classified as positive, but actually are not required objects (False). Looking at all images that were used in the training process, we get that the algorithm has successfully classified 705 objects from total number of 1000 used objects (70.5%). Algorithm did not recognize (29.5%) the remaining 295 objects. In addition, the algorithm is recognized 7600 objects as requested objects, though they are not. This means that the algorithm, by one successfully recognized object, averagely recognizes seven to eight objects as positive. These results can be characterized as expected when compared with other researches that have used the AdaBoost algorithm. Baluja and Rowley have used the same algorithm, trying to recognize gender based on the images that contain people faces [24]. The training gave an accuracy of 80%. This is somewhat higher accuracy, but the images used in testing using only the face which implicates minor amount of noise. When the training sets contain images with variable background, the accuracy of training process is decreasing. Yuan et al. showed differences in accuracy in recognizing faces when the background changes [25]. When the images contained different amounts of light, the resulting accuracy was 68.4%. In recognition of computer-generated characters with a constant background, AdaBoost achieves an accuracy of 90.5% [26].

The complete output obtained after testing can be presented using ROC curves as given in Fig. 2. This curve shows recognition we can expect when we allow a certain degree of false positives. Figure shows that when we allow seven or more false positive, the level of hits is about 70%. By reducing allowed number of false positive, the percent of true positive is reduced as well. This decrease is approximately linear up to the value 2 for false positives, where the percentage of correct hits is just below 55%. By further reducing the level of false positive, the level of true positive decreases exponentially and if we do not allow false positive, i.e. how this value approaches zero, we can expect only about 25% of successful recognition.

VI. Conclusion

Application of computer vision in the analysis of sporting events is quite common practice, especially in recent years. Basketball, as one of the most popular sports in which large amount of money is invested, does not deviate from this trend.

Three approaches can be applied in analyzing basketball games: an analysis using markers that are placed on basketball players, the analysis using multiple synchronized cameras that cover the whole court and image analysis using only one camera. This paper reports the third type of analysis, i.e. analysis of games using the videos broadcasted by television stations. This type of analysis brings the greatest amount of assumptions and inaccuracies, because players are often obscured by other players, or are outside the current view field of active camera.

Procedure to identify the players is trained by Gentle AdaBoost algorithm, which is trained on two sets of examples. The first set of images represented entire basketball players (head, body, arms and legs), while another set of images representing basketball players' upper body (head and body without arms or legs). AdaBoost failed to create a classifier based on images from the first training set. The reason is the large difference in the examples for training. The appearance of basketball players varies greatly depending on whether they walk, run, play defense, dribble the ball or shoot to the basket. In contrast, another training set, which included only head and body of basketball players, contained a much smaller degree of variation. This resulted in successfully finished training process by AdaBoost algorithm, which produced a classifier that can be applied in basketball player recognition. This algorithm has demonstrated an accuracy of 70.5%, which is not applicable in commercial applications. In addition, the algorithm has a number of areas in the pictures marked as basketball players (false positive).

This paper presented a degree of applicability of AdaBoost algorithm in the recognition of baseball players, without any prior processing. In order to improve obtained performances, we could apply background subtraction techniques that would leave only the objects that are likely to represent basketball players, on which the algorithm would then be applied. Further improvement would be achieved by mapping areas of interest, i.e. play field in observed application. This would remove everything that is not on the play field, which would make the search faster and more accurate. Another possible improvement would be training of AdaBoost algorithm for body parts (head, legs, arms, torso), which can then be combined in order to identify players.

ACKNOWLEDGMENT

Research was partially supported by the Ministry of Science and Technological Development of Republic of Serbia, through project Grant No: 171039.

REFERENCES

[1] T. Moeslund, A. Hilton, and V. Kruger "A Survey of Advances in Vision-based Human Motion Capture and Analysis", *Computer Vision and Image Understanding*, vol. 104, pp. 90-126, 2006.

[2] R. Schumaker, O. Soliman, and H. Chen, *Sports data mining*, Springer, 2010.

[3] B. Markoski, P. Vasiljevic, B. Petrevski, D. Radosav, Z. Ivankovic and Z. Milosevic "Basic statistical reports in basketball", *YUINFO*, Kopaonik, Serbia, 2010.

[4] B. Markoski, Z. Ivankovic, M. Nikolic and S. Hotomski, "Difficulties in basketball statistics", *Techniques, Informatics and Education TIO*, Cacak, Serbia, 2010.

[5] C. Ballard, "Measure of success", *Sports Illustrated*, 2005.

[6] Z. Ivankovic, M. Rackovic, B. Markoski, D. Radosav and M. Ivkovic, "Appliance of Neural Networks in Basketball Scouting", *Acta Polytechnica Hungarica*, vol. 7, no. 4, pp. 167-180, 2010.

[7] C. Barron and I.A. Kakadiaris "Estimating Anthropometry and Pose From a Single Image", *Computer Vision and Pattern Recognition*, 2000.

[8] C.J. Taylor "Reconstruction of articulated objects from point correspondences in a single image", *Computer Vision and Image Understanding*, vol. 80, no. 3, pp. 349-363, 2000.

[9] V. Parameswaran and R. Chellappa, "View independent human body pose estimation from a single perspective image", *Computer Vision and Pattern Recognition*, 2004.

[10] N. Krahnstoever and R. Sharma "Articulated models from video", *Computer Vision and Pattern Recognition*, 2004.

[11] Y. Song, I. Goncalves and P. Perona, "Unsupervised learning of human motion", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 7, pp. 814-827, 2003

[12] G. Cheung, S. Baker and T. Kanade, "Shape-from-silhouette for articulated objects and its use for human body kinematics estimation and motion capture", *Computer Vision and Pattern Recognition*, 2003.

[13] C. Menier, E. Boyer and B. Raffin, "3D skeleton-based body pose recovery", *International Symposium on 3D Data Processing, Visualization and Transmission*, 2006.

[14] G.J. Brostow, I. Essa,D. Steedly and V. Kwarta, "Novel skeletal representation for articulated creatures", *European Conference on Computer Vision*, 2004.

[15] J. Mulligan, "Upper body pose estimation from stereo and hand-face tracking", *Canadian Conference on Computer and Robot Vision*, 2005.

[16] L. Herda, R. Urtasun and P. Fua, "Hierarchical implicit surface joint limits for human body tracking", *Computer Vision and Image Understanding*, vol. 99, no. 2, pp. 189-209, 2005.

[17] T.B. Moeslund, C.B. Madsen and E. Granum, "Modeling the 3D pose of a human arm and the shoulder complex utilizing only two parameters", *International Journal on Integrated Computer-Aided Engineering*, vol. 12, no. 2, pp. 159-177, 2005.

[18] A. Micilotta, E. Ong and R. Bowden, "Detection and tracking of humans by probabilistic body part assembly", *British Machine Vision Conference*, Oxford, UK, 2005.

[19] Y. Yang and D. Ramanan, "Articulated Pose Estimation using Flexible Mixture of Parts," in *Computer Vision and Pattern Recognition (CVPR)*, pp. 1385-1392, 2011.

[20] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", *Computer Vision and Pattern Recognition*, 2001.

[21] Y. Freund and R.E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119-139, 1997.

[22] A. Kuranov, R. Lienhart and V. Pisarevsky "An empirical analysis of boosting algorithms for rapid objects with an extended set of haar-like features", *Intel Technical Report MRL-TR*, 2002

[23] P. Viola, M. Jones and D. Snow, "Detecting pedestrians using patterns of motion and appearance", *International Journal of Computer Vision*, vol. 63, no. 2, pp. 153-161, 2005.

[24] S. Baluja, and H. Rowley, "Boosting Sex Identification Performance", *International Journal of Computer Vision*, vol. 71, no. 1, pp. 111-119, 2007.

[25] Z. Yuan, Z. Lu and H. Pan, "LPP-AdaBoost based face detection in complex backgrounds", *Fourth International Conference on Computer Sciences and Convergence Information Technology*, pp. 451-456, 2009.

[26] L. Li, K.E. Hoe, X. Yu, L. Dong and X. Chu "Human Upper Body Pose Recognition Using Adaboost template for natural human robot interaction", *Canadian Conference Computer and Robot Vision*, pp. 370-377, 2010.