Computer Vision Winter Workshop 2007, Michael Grabner, Helmut Grabner (eds.)
St. Lambrecht, Austria, February 6–8
Graz Technical University

Automatic Evaluation of Organized Basketball Activity using Bayesian Networks

Matej Perše, Matej Kristan, Janez Perš, and Stanislav Kovačič

Faculty of Electrical Engineering, University of Ljubljana,
Tržaška 25, SI-1000 Ljubljana, Slovenia
matej.perse@fe.uni-lj.si

Abstract *In this article the trajectory-based evaluation of multi-player basketball activity is addressed. The organized basketball activity consists of a set of key elements and their temporal relations. The activity evaluation is performed by analyzing individually each of them and the final reasoning about the activity is achieved using the Bayesian network. The network structure is obtained automatically from the activity template which is a standard tool used by the basketball experts. The experimental results suggest that our approach can successfully evaluate the quality of the observed activity.*

1 Introduction

The objective analysis of player performance is one of the main goals in the field of sport science. While individual player's physical capabilities can be readily tested in laboratory conditions, the team performance can only be observed during the actual gameplay. This process may include an advanced statistical analysis based on video recording, but it nevertheless relies on observation and manual annotation by sport experts, with the potential risk of becoming too subjective. Moreover, manual annotation is a time consuming and tedious task, mostly limited either to the academic research or to the teams which can afford a sufficient number of qualified experts.

This is a reason for an increasing volume of research towards automatic or semi-automatic trajectory-based analysis of human behavior in sports. The ultimate goal of such research is to develop methods for automatic interpretation and analysis of team performance, which would present a concise summary of the team's and players' strengths, weaknesses and mistakes. In addition, the same methods could be used in many other areas, such as sports broadcasting, either for the purpose of enhancing live broadcasts or facilitating video archival. In the broader context, similar methodology could be used in human motion analysis for video surveillance, ambient assisted living, and related areas. Nevertheless, the main focus of this article is the challenge of analyzing the organized team activity on the field.

This paper is structured as follows. The rest of this section describes the related work and the motivation behind our approach. Section 2 presents the structure of the basketball game and properties of the basketball templates. Section 3 describes how the Bayesian network is constructed from a template. Section 4 presents the recognition results and other advantages of our approach. Finally, in Section 5 we discuss the results and give the guidelines for the future work.

1.1 Related work

The analysis of the trajectory-based multi-agent behavior is an active research domain since the trajectories contain useful information about the agents' behavior. Various approaches dealing with trajectory-based recognition and reasoning have been presented in the last decade. For example, Hidden Markov Models (HMMs) [14, 8], Fuzzy logic [9], Finite State Machines [12] or Bayesian networks [6], were used for modeling the basic logical and temporal relations of events which are of key importance for the understanding of the overall behavior.

Although HMMs are a powerful tool for modelling temporal relations, they have proved to be less appropriate for domains in which multi-agent relationships result in large feature spaces [5]. Intille and Bobick [6] propose multi-agent Bayesian (belief) network which reflects the temporal structure of the activity and uses the temporal analysis functions and local visual networks for the observable evidence nodes. In the course of the activity recognition these nodes provide the evidence about the occurrence of the basic events from which the likelihood of particular activity is calculated. This approach is reasonably successful for recognizing the type of the activity the team is performing. However, their visual networks and temporal functions have to be constructed manually for every single activity event. This demands a lot of engineering background and is therefore impractical for the everyday use by sport experts.

1.2 Our approach

This paper extends the original idea of multi-agent activity recognition [6] in a way that it can be used by the sport experts on a daily basis. Similar to [6], we use the multi-agent belief networks, detectors of basic events and temporal relation functions, however, in our case the structure of the network and the temporal relations are *automatically obtained from the activity template* which sport experts generally use when preparing the tactics of the game [3, 19, 10]. In contrast to [6], where authors use very complex visual networks to represent the basic activity elements, we use just two simple trajectory-based detectors of basic events (*move* and *screen* detectors described in the next section) and three temporal relations (*before*, *within* and *around*). These detectors proved to be sufficient for the evaluation of the team performance.

2 Structure of the basketball game

From the analytical point of view, the basketball game can be divided into different game phases, such as offensive game, defensive game, active or inactive game, etc. These phases can be additionally divided into game sub-phases according to the type of the phase. For example, the offensive game can be subdivided into the organized or unorganized offense and counterattack, and the defensive phase can be subdivided into zone defense, men-to-men defense and various combinations of the two.

This work focuses only on the analysis of the organized basketball offense which is the most important segment of the offensive game phase for basketball experts. This type of offense can be described as an activity comprised of independent basic basketball elements (e.g. *player motion, dribbling, passing, shooting, screening, rebounding, team starting formation*, etc.), which have to be executed in a particular temporal order [10]. When designing such activities, sport experts often use basketball templates [3, 19, 10]. Our implementation of basketball template specifies the spatio-temporal properties of each individual element, i.e. the position where the element should occur and the temporal interval in which it should occur.

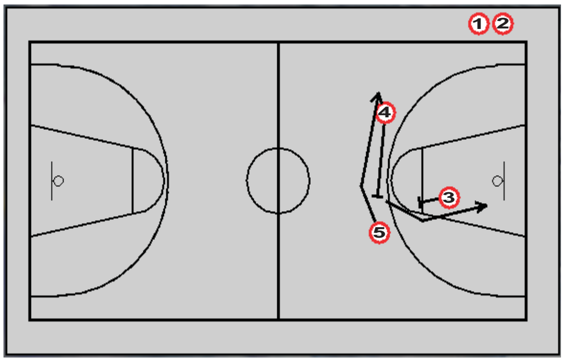
Figure (1) shows an example of the template for a simple basketball action called ”Double Screen”. This activity is comprised of four player *moves* (player movements along predefined path) and two *screens* (close contacts of two players, where one of the players is standing still) and could be described as follows:

- Player 4 should first move to the position of the screen.
- After player 4 has positioned himself to the position of the screen, player 5 should run next to player 4 and use the screen. At the same time, player 3 should position himself to the position of screen for player 4.
- Finally, when players 3 and 5 have moved to their new positions, player 4 should move next to player 3 and use the screen.

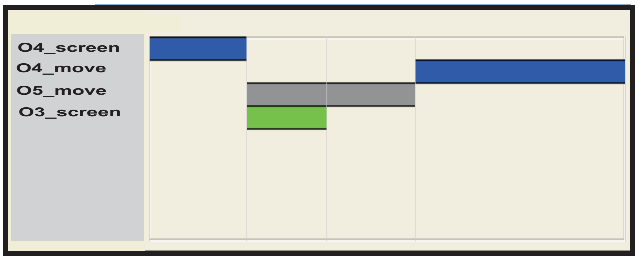
The above description contains all the information that is important for the correct execution of offensive action. Therefore, by evaluating how well have the elements been performed and what where the temporal relations among them, we could establish the overall score of the observed activity.

2.1 Evaluation of basic basketball elements

In general, the basketball elements can be grouped according to the ball possession to those elements that involve the ball (*dribbling, passing, shooting, rebounding*) and the elements that do not involve the ball (*player motion, screening, team starting formation*). Since our analysis is based solely



(a)



(b)

Figure 1: An example of spatial (a) and temporal (b) properties of a template which represents the organized offensive activity called ”Double Screen”.

on players’ trajectories, we are only able to analyze the elements that do not involve the ball[1]. Therefore we determine the models of the detectors for the following two basketball elements (*screen* and *move*):

- ***Screen***. In the basketball literature, the *screen* is defined as a close contact between two players [10], where ideally one player is standing still and the other runs in his near proximity. Thus a certain interaction among two players is more likely to be interpreted as a *screen* if the velocity of the slower player is low and the distance between the players is small.

  Let d_t be the Euclidian (l_2) distance between the two interacting players and let v_t be the velocity of the slower player. The likelihood function of a *screen* is defined as

$$\mathcal{L}(screen|d_t, v_t) \stackrel{\Delta}{=} \mathcal{N}(d_t; 0, \sigma_d) \cdot \mathcal{N}(v_t; 0, \sigma_v), \quad (1)$$

  where $\mathcal{N}(\cdot; 0, \sigma)$ is a zero-mean Gaussian function with variance σ^2. We define the quality of the *screen* as the likelihood ratio

$$S_{screen} \stackrel{\Delta}{=} \frac{\mathcal{L}(screen|d_t, v_t)}{\mathcal{L}(screen|0, 0)}, \quad (2)$$

  where $\mathcal{L}(screen|d_t, v_t)$ is the likelihood of the *screen* given the current distance and velocity values (d_t, v_t) of the interacting players and $\mathcal{L}(screen|0, 0)$ is the likelihood of an ideal *screen*.

[1]Some of the elements such as dribbling and rebounding are also observable, but since we are not able to determine if the ball is actually involved we regard them as a basic *player motion*.

The proximity parameter σ_d in (1) is set to $\sigma_d = 1$m. The velocity parameter σ_v which determines the velocity of player which is "*still enough*" is set to $\sigma_v = 0.5$m/s in our implementation.

- ***Move***. In the activity template, a player *move* is defined as the exact path that a player should follow. The path is defined by one or more line segments where each segment has a starting point (A_i) and an ending point (B_i). Therefore, we can define the quality of the player's move as a product of the distance function from the ideal path $\mathcal{N}(d_t; 0, \sigma_d)$ and the path ratio $f_{path}(t)$

$$S_{move}(t) \triangleq \mathcal{N}(d_t; 0, \sigma_d) \cdot f_{path}(t). \quad (3)$$

Parameter d_t in the equation (3) denotes the l_2 distance between player and the closest point on the path (perpendicular distance) and function $f_{path}(t)$ determines the ratio between the path that player has covered up to time t and the total path

$$f_{path}(t) = \frac{\sum\limits_{i=1}^{M} \sum\limits_{j=1}^{t} (\Delta\overrightarrow{x_j} \cdot \overrightarrow{A_i B_i})}{\sum\limits_{i=1}^{M} ||\overrightarrow{A_i B_i}||}, \quad (4)$$

where M is the number of path segments, $||\overrightarrow{A_i B_i}||$ defines the length of path segment i and $\sum\limits_{j=1}^{t} (\Delta\overrightarrow{x_j} \cdot \overrightarrow{A_i B_i})$ defines the sum of the scalar products of current player motion vector $\Delta\overrightarrow{x_j}$ and ideal motion vector of i^{th} segment $\overrightarrow{A_i B_i}$.

Due to the nature of the probabilistic reasoning model (Bayesian network) which is used to reason about the activity score, the maximal values of detector outputs are used as the evidence, which is entered to the evidence nodes of the reasoning network. The reason for this is that if the current detector outputs were used for the evaluation, the final estimate of the activity score would be low since the elements occur in a certain temporal order (not all at the same time) and detectors outputs return to zero after the elements are performed. However, the temporal profile of the detector output is considered in the final estimate through the temporal relation functions which determine the temporal relations among elements.

2.2 Evaluation of temporal relations

Temporal relations (TR) define the order in which players should perform the individual basic elements. Therefore we define three different temporal relations (*before*, *within* and *simultaneously*) between the two elements (*screen* and *move*):

- Element $E1$ should occur ***before*** element $E2$. To determine this temporal relation, we first calculate the temporal gravity centers t_{E1} and t_{E2} for detectors $E1$ and $E2$, respectively according to the equation

$$t_X = \frac{\sum\limits_{i=1}^{N} t_i \cdot S_X(i)}{\sum\limits_{i=1}^{N} t_i}, \quad (5)$$

where $X \in \{E1, E2\}$, N stands for the number of observations and $S_X(i)$ is the detector output at time t_i. The quality of the relation *before* is defined as

$$TR_{before} = \begin{cases} 0 & t_{E1} > t_{E2} \\ k \cdot g(t_{E2} - t_{E1}; \mu, \sigma) & t_{E1} <= t_{E2} \end{cases}, \quad (6)$$

where $g(x; \mu, \sigma)$ is the log-normal function

$$g(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\{\ln x - \mu\}^2/2\sigma^2}, \quad (7)$$

with $\mu = 4s$ and $\sigma = \sqrt{2}s$ and k is the normalization constant such that $g(x; \mu, \sigma) \in [0, 1]$.

The log-normal distribution function is used because we want to additionally penalize temporal intervals Δt that are too short (i.e. shorter than half a second). We assume, that the ideal temporal interval has a width of approximately $1s$.

- Element $E1$ should happen ***within*** element $E2$. This temporal relation represents two events, where one of the two starts and ends inside the other. The score of this relation is defined as

$$TR_{within} = \frac{\sum\limits_{i=1}^{N} min\{S_{E1}(i), S_{E2}(i)\}}{min\{\sum\limits_{i=1}^{N} S_{E1}(i), \sum\limits_{i=1}^{N} S_{E2}(i)\}}. \quad (8)$$

Function (8) is an area-based similarity measure of the two functions. This similarity is expressed as the ratio between the cross-section of detector output and the area of the smaller output which is supposed to happen within.

- Element $E1$ should happen ***simultaneously*** with element $E2$. This temporal relation represents two events that should start and end approximately at the same time. It is defined as

$$TR_{sim} = \frac{\sum\limits_{i=1}^{N} min\{S_{E1}(i), S_{E2}(i)\}}{\sum\limits_{i=1}^{N} max\{S_{E1}(i), S_{E2}(i)\}}. \quad (9)$$

This function defines the ratio between the cross-section and union of the observed outputs. In case when the outputs are perfectly aligned the cross-section equals the union and therefore the result of the score function equals one.

These three temporal relation functions represent minimal set of functions that are needed to represent the temporal relations between the basketball elements.

3 Building the Bayesian network

A Bayesian network is a directed acyclic graph that encodes a joint probability distribution over a set of random variables $U = \{X_1, ..., X_N\}$ where each variable X_i takes on values $x_1, ..., x_M$. Formally, the Bayesian network for U is a pair $B = \{G, \Theta\}$, where G denotes the directed acyclic graph whose nodes represent the variables and arcs represent the dependencies between the variables. The second component Θ, is a set of parameters of the network, which represents the conditional probabilities $P(X_i|Pa\{X_i\})$ for all parents $Pa\{X_i\}$ of variable X_i. Given B we can calculate the joint probability distribution over U using the so called *chain rule*

$$P(U) = \prod_{i=1}^{N} P(X_i|Pa\{X_i\}). \tag{10}$$

Given the joint probability $P(U)$, we can calculate the a priori probability $P(X_i)$ for any variable X_i from U by marginalizing the joint distribution $P(U)$ over all other variables from U (see [16, 7] for details). However the main purpose for building the network is to use it for inference. In our case, we want to use the network for *belief updating*, which can be described as the calculation of posterior probability $P(X_i|e)$ of node X_i given the evidence e which represents the values of evidence nodes.

In our case, the Bayesian network is used to calculate the posteriori probability that a certain activity has been executed, given the evidences about the execution of basic elements and temporal relations. The execution of the basic elements is obtained by applying the element detectors to the player trajectories. The set of detectors could be viewed as a set of independent referees where every referee has a task to asses an individual element and the Bayesian network could be interpreted as the agreement of all the referees about the final activity score and the validation of individual player performances.

This article focuses mainly on the structure of the network G since, as experiments suggest, it has a greater impact on the analysis results than the conditional probabilities Θ. The structure of the network specifies the type of key elements and the number of key temporal relations among them. In general, this could be done from training data or with the help of the basketball expert.

Several approaches concerned with the construction of the Bayesian network from the training data have been proposed in the last decade [15, 4, 1]. The main problem with these approaches is that they demand large amounts of training data for building networks with relatively small number of attributes. For example Cheng [2] used one thousand test samples for building the network consisting of eight attributes and more than 10.000 samples for building the network with 37 attributes. In our case this would mean, that the same activity should have been repeated at least one thousand times in order that this type of network construction would be possible. This would take too much time and is perhaps acceptable for research purposes but is totaly impractical for the everyday use.

Another option is to construct the Bayesian network with the involvement of a basketball expert. In this case the expert's task is to provide all the information about the relations among variables. The main drawback of this approach is that the sport expert has to be familiar with the probabilistic theory and the characteristics of the Bayesian networks or it has to collaborate with a computer expert who is familiar with the previously mentioned concepts. In later case the computer expert could guide the sport expert through the modeling procedure by using specific net-construction rules such as the ”Causal mapping approach” presented by [13]. Both approaches are time consuming and expensive. This motivated us to device a method which would automatically construct a Bayesian network from the existing activity template.

3.1 Building the Bayesian network from activity template

The activity template already contains all the information needed to construct the network. It specifies the number and type of elements and their temporal profile. This temporal profile can be automatically transformed into the temporal relations using some heuristic rules.

In order to be able to define the temporal relations between elements, we generate synthetic trajectories using the spatio-temporal information from the template and apply the element detectors to those trajectories. This way we obtain the *activity timeline* that defines the actual time intervals in which the elements occur. It should be noted that the *screen* element, as it is defined in the template, is divided into the *move* of the player that is making the *screen* to the screen position and the actual *screen* element (e.g. elements *pl 4 move* and afterwards *pl 4 screen for pl 5* in Figure 2).

By observing the starting and ending times of the element intervals, we can define whether the element has to be executed *before*, *within* or *simultaneously* according to other elements from the activity. It is important, that we retain only the relations to those elements that happen immediately before the element that is currently analyzed. This approach does not guarantee the optimal definition of temporal relations, since it may happen that some additional relations between elements which are not contextually dependent are observed (e.g. *pl 4 screen for pl 5* and *pl 3 move* in Figure 2). However, such temporal relations should be observed anyway whenever the activity is executed correctly and thus should not influence considerably the final activity evaluation.

Once the temporal relations are established, the network building procedure can begin. The Bayesian network is divided into four contextual levels (Figure 3):

- Level four (the lowest level) nodes represent the temporal relations between different elements. Their probabilities are obtained from the outputs of temporal relation functions.
- Level three nodes represent the basic elements. They carry the information about the execution of each individual element. Their probabilities are obtained from the outputs of the element detectors.
- Level two contains the *player nodes*. These nodes define the performances of individual players.

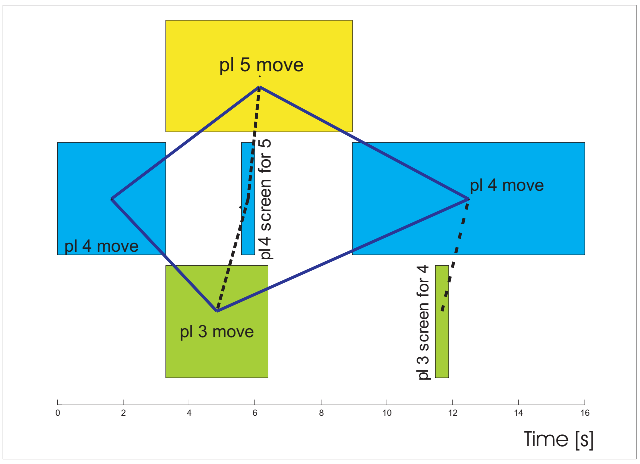


Figure 2: The timeline for the *Double screen* activity. The lines represent he learned temporal relations. Full lines represent relation *before* and dashed lines represent relation *within*.

- Level one (the highest level) contains the *activity node*. This node defines the overall execution correctness of the observed activity.

  The probabilities of nodes on the first and the second level can not be observed directly from the trajectories and therefore depend on the evidence from nodes on levels three and four and are given as posterior probabilities $P(A_i = true|e)$. The higher the posterior probability the greater the chance that the activity was performed flawlessly.

Figure 3 shows an example of Bayesian network for the ”Double screen” template shown in Figure 1.

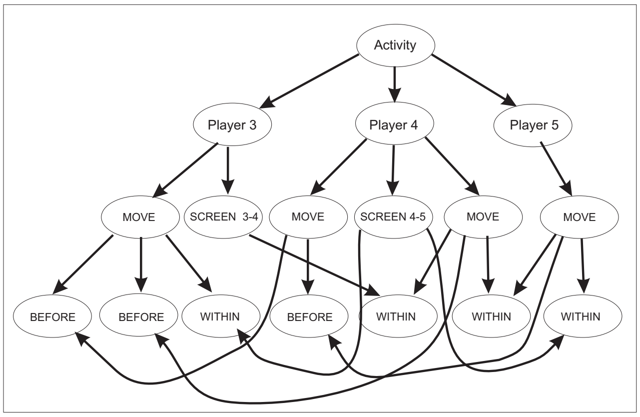


Figure 3: Example for the Bayesian network for the *Double screen* activity. Nodes represent the variables and the arrows represent probabilistic relations between variables.

The obtained structure contains all the key elements which are linked with the correct temporal relations. However, from the template we are only able to obtain the structure of the network, but in order to use the network we also need to derive the conditional probabilities. This problem has not yet been fully addressed, therefore we currently define these probabilities so that the *a priori* probabilities of all the node variables are equal for all nodes ($P(X_i = true) = 0.5$ and $P(X_i = false) = 0.5$) which means that the network does not contain any a priori knowledge. We tested different *a priori* probabilities which yielded similar results. However, we believe that even better *a priori* and conditional probability estimates could be obtained by analyzing the outputs of different detectors or by considering the importance of element that a player has to perform. This information is also encoded in the activity template, however it is currently not considered.

4 Experimental results

To test the performance of our activity evaluation method, we acquired 63 trajectory segments of three different types of basketball activities. The first type of activity was repeated 21 times, the second 24 and the third 18 times. To obtain the trajectory segments, we performed supervised tracking using a probabilistic color-based tracking algorithm [11, 17]. Since the evaluation procedure requires that the roles of players are known (i.e. we have to know which trajectory represents which player role in the template), we cast players into their respective roles using the method described in [18].

Our first goal was to determine if it is possible to establish the type of the activity which the team is performing solely from the final activity score. To do that, we have built the appropriate template for each of the three basketball plays and 14 additional templates, obtained from basketball literature [19, 10]. The templates have been transformed into Bayesian networks with the procedure described in Section 3.

Table 1 shows the confusion matrix of the average activity scores when different templates are used to build the network. The average activity score was calculated as the mean value of all the scores for a particular type of activity and represents the score that an average team would receive when playing the activity in an average manner (i.e. it includes some perfectly executed activities as well as some of those with several errors). However, since the average score is not useful in cases when we want to determine the threshold that separates the positive and negative examples, Table 1 also shows the maximum posterior activity score for the cases when the type of template and test example did not match and the minimum posterior score for examples when the type of template and test examples matched.

As it can be observed from Table 1, it is possible to classify the observed activity solely by the score it receives in the evaluation procedure, since the average score is at least three times greater in cases when the types of Bayesian network and the performed activity match. The results for the minimum match case and maximum mismatch case also confirm the above hypothesis and suggest that a suitable threshold which separates the positive and negative examples could be set at a score value of 60%. From this we can conclude that Bayesian network, constructed automatically from the activity template can be used to recognize the type of the play which the team is performing.

The second important type information that can be obtained from the network is the performance of individual

Table 1: The average efficiency score given as the posterior probability $P(act = true|e)$. The best evaluation result obtained for each activity type is displayed in bold. The ideal activity would have the score of 100 %. The bottom two rows show the maximum score of all examples when the template and test example mismatch and the minimum of all cases when tested example match.

	Activity scores [%]		
Template type	$P(Off1\|e)$	$P(Off2\|e)$	$P(Off3\|e)$
T1	**89.06**	16.99	4.51
T2	10.66	**87.08**	3.46
T3	2.42	1.96	**90.52**
T4	11.94	32.51	7.86
T5	5.96	2.71	2.64
T6	5.51	11.20	16.76
T7	2.99	3.52	21.59
T8	16.48	15.25	13.50
T9	6.91	7.95	16.78
T10	3.43	8.61	4.07
T11	2.29	3.15	4.41
T12	8.71	28.44	4.14
T13	10.98	6.16	5.55
T14	5.07	3.36	3.11
T15	18.47	11.30	2.23
T16	13.01	9.22	5.22
T17	19.25	9.55	3.88
Min. match	76.38	72.31	74.08
Max. mismatch	33.78	53.93	37.59

players and the reasons for the good or poor performance of the team. This information can point coaches and players to the mistakes the players are making when performing the activity since each element is graded separately and this way can help them to improve the individual elements and the overall performance.

Figure 4 shows the examples of two networks for the same type of activity. Example (a) represents a well executed activity and example (b) represent an activity which was executed poorly. By studying in detail the two networks, we can establish, that the main reason for the poor execution of activity in Figure 4 (b) lies in the poor performance of players 2 and 5 (p2 and p5). Furthermore, we can observe, that the main cause for their low grades is the lack of execution of the *screen* (*screen* 5-2). Additionally player 2 also failed to perform the *screen* with player 1 (*screen* 1-2) and therefore received the lowest score. The same conclusion was obtained by studying the archived activity video, which showed that player 2 performed a wrong move in the middle of the activity and therefore failed to execute the elements that should follow.

The immediate result of the above example study could be the suggestion to the basketball coach that he should substitute players 2 and 5 or that the players should additionally train the activity with the special focus on the elements which obtained the lowest scores.

5 Conclusion and future work

An approach for automatic evaluation of basketball activities with use of Bayesian networks was presented. The networks were built automatically from the activity templates which are commonly used by sport experts to pass their ideas to players. The basic network structure was divided into four levels. The first and the second level represent the overall activity score and the scores of individual players, respectively. The values of the variables in these two levels depend directly on the values of the lower two levels, which represent the execution of the key activity events (level three) and their temporal relations (level four). The values for the lower two levels are given in the form of soft evidence and are obtained by the trajectory-based element detectors and temporal relation functions.

Based on the results of testing on 63 real trajectory segments that vary significantly in time and space, we can conclude that Bayesian network which is obtained from the activity template can be used for the recognition of the type of the performed activity. The results demonstrated that the final activity score is three times greater in cases when type of activity and type of the network match. Additionally, the insight into the graphical structure of the Bayesian network is helpful in further analysis the overall performance, i.e. when we want to determine why a particular activity was performed poorly. This can help the basketball experts to discover and eliminate the causes for poor performance, thus giving the team an additional insight into their game.

The presented approach could be extended to work in other scenarios where the behavior can be specified by (possibly) multiple activity templates. For example, in surveillance such approach could be used to differentiate between the expected and the unexpected pedestrian behavior.

Future work should be devoted to the analysis of the influence of conditional probabilities to the evaluation results. These probabilities could be determined in such way that they would reflect the importance of individual elements for the overall action performance. Additionally, the evaluation results should be validated by the basketball experts in order to establish if the analysis results really reflect the opinion of experts.

Acknowledgement

This research was supported in part by Slovenian Research Agency (ARRS) contracts P2-0232 and L5-6274, and Slovenian Ministry of Defence (MORS) contract CiVaBis M2-0156.

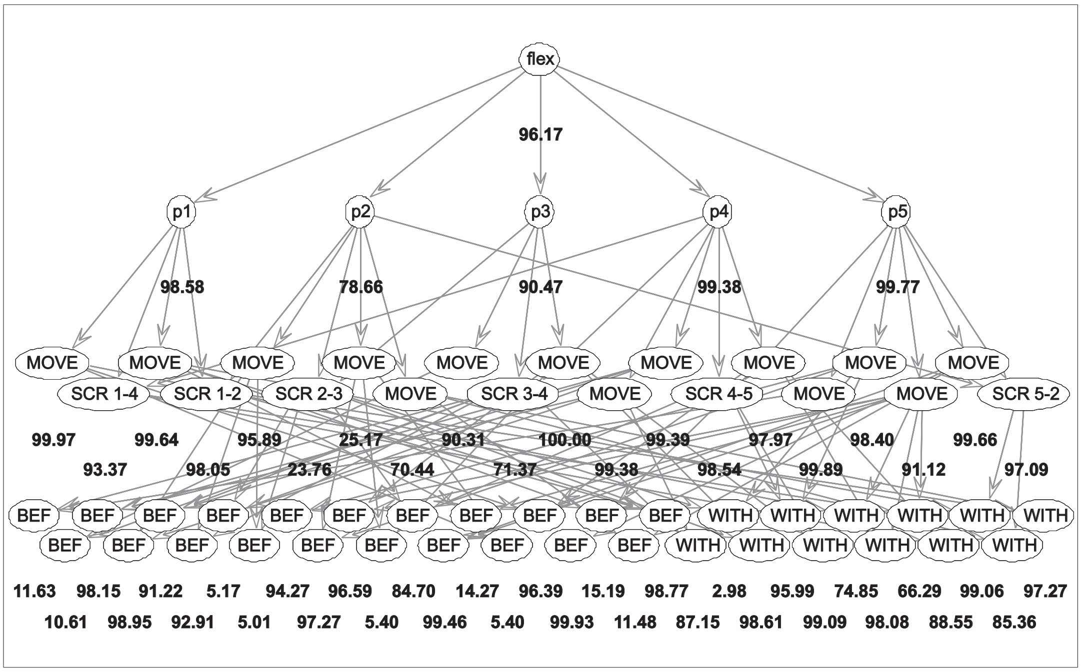
References

[1] Wray L. Buntine. Operations for learning with graphical models. *Journal of Artificial Intelligence Research*, 2:159–225, 1994.

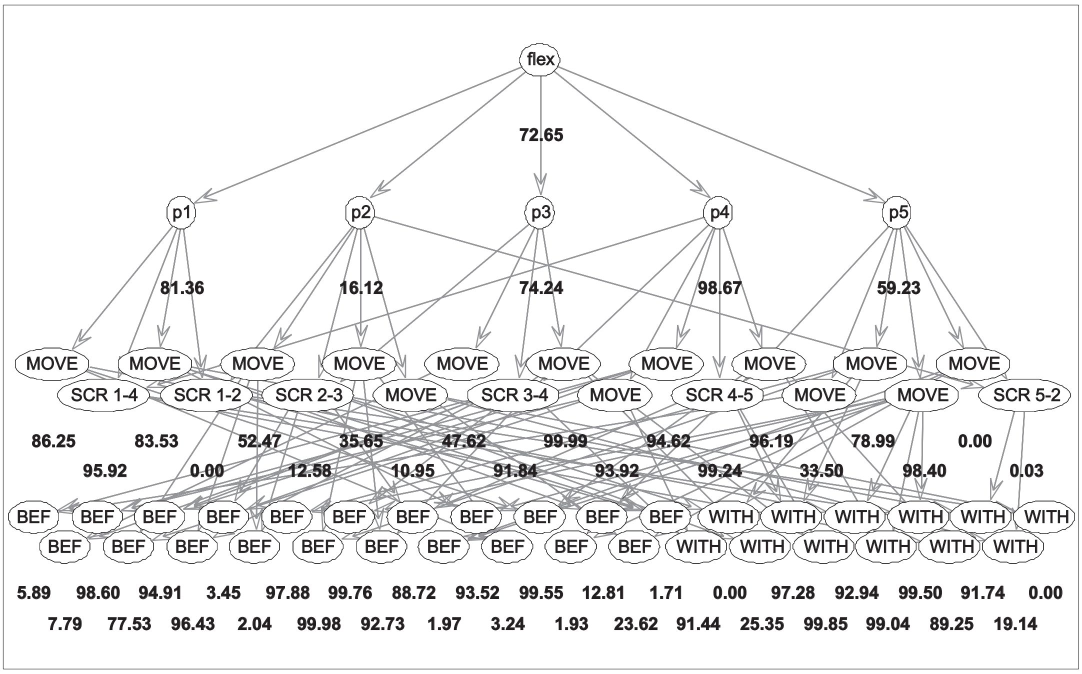
[2] Jie Cheng, David A. Bell, and Weiru Liu. Learning belief networks from data: An information theory based approach. In *CIKM*, pages 325–331, 1997.

[3] H.L. Harkins and J. Krause. *Motion Game Offenses - For Men's and Women's Basketball*. Coaches Choice, 3 edition, 2001.

[4] D. Heckerman. A tutorial on learning with bayesian networks. Technical report, A tutorial on learning with bayesian networks. Technical Report



(a)



(b)

Figure 4: Examples of scores for two activities of the same type. The numbers below the nodes represent the posterior scores for each node. (a) Well executed activity. (b) Badly executed activity. The network shows that the player 2 and 5 performed their elements poorly.

MSR-TR-95-06, Microsoft Research, Redmond, Washington, 1995.

[5] S.S. Intille and A.F. Bobick. A framework for recognizing multi-agent action from visual evidence. In *in: Proceedings of the sixteenth national conference on Artificial intelligence and the eleventh Innovative applications of artificial intelligence conference innovative applications of artificial intelligence, Orlando, Florida, United States*, pages 518–525, July 1999.

[6] S.S. Intille and A.F. Bobick. Recognizing planned, multiperson action. *Computer Vision and Image Understanding: CVIU*, 81(3):414–445, 2001.

[7] F.V. Jensen. *An introduction to Bayesian networks*. UCL Press, London, 1996.

[8] H. Johnson and D. Hogg. Represention and synthesis of behavior using gaussian mixtures. *Image and Vision Computing*, 20:889–894, 2002.

[9] M. Jug, J. Perš, B. Dežman, and S. Kovačič. Trajectory based assessment of coordinated human activity. In *Computer Vision Systems, Proceedings of Third International Conference ICVS 2003*, pages 534–543, Graz, Austria, April 2003.

[10] J. Kresse and R. Jablonski. *The Complete Book of Man-To-Man Offense*. Coaches Choice, 2nd edition, 2004.

[11] M. Kristan, J. Perš, M. Perše, M. Bon, and S. Kovačič. Multiple interacting targets tracking with application to team sports. In *4th International Symposium on Image and Signal Processing and Analysis ISPA*, pages 322–327, September 2005.

[12] F. Li and R.J. Woodham. Analysis of player actions in selected hockey game situations. In *Proceedings of the The 2nd Canadian Conference on Computer and Robot Vision (CRV'05), Canada*, pages 152–159, 2005.

[13] S. Nadkarni and P.P. Shenoy. A causal mapping approach to constructing bayesian networks. *Decision Support Systems*, 38:259–281, 2004.

[14] V. Nair and J. Clark. Automated visual surveillance using hidden markov models. In *VI02 - International Conference on Vision Interface*, pages 88–93, 2002.

[15] R.E. Neapolitan. *Learning Bayesian Networks*. Prentice Hall, 1 edition, April 2003.

[16] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: networks of plausible inference*. Morgan Kaufmann, 2 edition, 1988.

[17] M. Perše, M. Kristan, J. Perš, M. G. Vučkovič, and S. Kovačič. Physics-based modelling of human motion using kalman filter and collision avoidance algorithm. In *International Symposium on Image and Signal Processing and Analysis, ISPA05, Zagreb, Croatia*, pages 328–333, September 2005.

[18] M. Perše, J. Perš, M. Kristan, and S. Kovačič. A template-based multi-player action recognition of the basketball game. In *In: Janez Pers, Derek R. Magee (eds.), Proceedings of the ECCV Workshop on Computer Vision Based Analysis in Sport Environments, Graz, Austria*, pages 71–82, May 2006.

[19] R. Righter. *Flex: The Total Offense*. Championship Books, Ames, Iowa, USA, 1984.