SOCCER PLAYER MOTION RECOGNITION BASED ON STATISTICAL WEIGHTS

Nam Sung Hyun
Doowon Technical College
678, Jangwon-ri, Juksan-myon, Ansung-shi, Kyonggi-do
Korea<456-718>
Tel.: 82-31- 670-7212/10
Fax : 82-31-670-7219
e-mail: shnam@doowon.ac.kr

Kim Hyun Sook
Shinsung College
San 49, Dukma-ri, Jungmi-myon, Dandjin-gun
Choongchungnam-do, Korea< 343-860>
Tel.: 82-41-350-1270
e-mail: khyuns@shinsung.ac.kr

Hwang Chong Sun
Korea University, Graduate School of Computer Science and Technology<136-701>
1, 5Ga Anam-dong Sungbook-gu, Souol, Korea
Tel.: 82-2-3290-4042
Fax : 82-2-927-9027
e-mail: hwang@disys.korea.ac.kr

Yang Young Kyu
Electronics and Telecommunications Research Institute
Image Processing Department<305-350>
161, Gajund-dong, Yoosung-gu, Daejun-shi, Korea
Tel.: 82-42-860-6380
Fax : 82-42-869-1479
e-mail: ykyang@etri.re.kr

ABSTRACT

In this work, we address a new approach to motion recognition of soccer players in a video scene. The inherent motion of a moving character's arms and legs and neighbouring position of the object, such as a ball, closely influences the motion of the soccer player.

In this paper, we choose the features for recognising a player's motion by human knowledge and describe correlation between objects. We propose an approach to motion recognition by statistical weights. Also we test, analyse, and estimate the proposed approach. Our algorithm is tried and tested on soccer games. We include successful results as well as a few examples of the algorithm's failures. Finally, we suggest some improvements and future extensions.

Keywords: ***Soccer, Motion recognition, Features, and Statistical weights***

1.0 INTRODUCTION

With rapid advances in data storage, image processing, telecommunications and data compression, video has become an important component in the new generation of information systems. Video data contain a large amount of spatial and temporal information. They can provide more information than text, graphics and image. The information related to the position, distance, temporal and spatial relationships are included in the video data implicitly. The use of image sequence is usually in the following areas: entertainment, visual communications, multimedia, education, medicine, motion recognition, scientific research, and sports.

Soccer produces many interesting fields in video analysis. It is very important to analyse objects, such as players and ball movements, etc., in soccer games. Soccer game analysis is largely classified into the following three parts: video scene change detection and indexing, multiple moving object extraction and presentation, and panoramic view synthesis. Multiple moving object extraction and presentation include motion recognition of moving objects and formation recognition in soccer games. Motion and formation recognition are very essential points in soccer game analysis.

In this paper, we recognise the motions of moving objects and we describe the approach to recognition in dynamic scenes. Fig. 1 shows the motions of soccer players. We choose the geometric features from a dynamic motion for recognition of moving objects and describe the correlation between objects. We propose the approach to recognition by statistical weights. Also we test, analyse, and estimate the proposed approach. The proposed approach is tested on soccer games. We introduce the study in Section 1.0, discuss related works and our work in Section 2.0, motions in Section 3.0, recognition of a player's motion in Section 4.0, experimental results in Section 5.0, and then we conclude and discuss the study in Section 6.0.

(a) Dribble (b) Run (c) Walk (d) Stand (e) Kick

Fig. 1: Motions of soccer players

2.0 RELATED WORKS AND OUR WORK

2.1 Related Works

Recently, computer vision techniques have been widely applied to sports games for automatic extraction and analysis of various information. Elisabeth Andre et. al. worked on the SOCCER system. They described automatic generation of simultaneous description calls for the application of an incremental event recognition strategy and for the adequate coordination of event recognition and language production. Their paper [1] is about the automatic generation of natural language comment or announcing during the soccer game. Also, Chueh-Wei and Suh-Yin Lee presented several important methods in a video information system for sport motion analysis. They described an effective calculation method for a moving object's positions on a video sequence, and described tracking objects and extraction features. They proposed a matching mechanism for comparing the difference between two video sequences. Their paper [2] is about moving object analysis, but not about motion recognition. Dennis Yow, et. al. presented techniques to automatically detect and extract soccer highlights by analysing the image contents, and presented these shots of action by analysis which includes the recognition of prominent features of the game. Their paper [3] discusses the research about tracking object and recovery.

Much research has been done on moving object analysis. Koh Kakusho et. al. discussed recognition of social dancing from a real image sequence and an acoustic signal of music. They tracked the power of the sound of music, the rotation and position of dancers to recognition for the kind of dance when dancers were dancing. Their paper [4] is not for the recognition of dancers' motion, but tracks the steps and positions of dancers, and recognises the kind of music. Lee Campbell and Aron Bobick have developed a system for recognising classical ballet steps from the input of Cartesian x,y,z tracking data. They developed techniques for representation of movements based on space curves in subspaces of a "phase space". They recognised nine movements of traditional ballet from two players. It is something similar to our recognition, but it is on the description of fixed position of an object and camera, and ballet movements are formulated [5]. Jia-Ching Cheng and Jos M. F. Moura recognised the motions of a person from a video sequence and proposed model-based recognition method on a dynamic environment. They discussed detecting a person and a position. They extracted and expressed modelling elements from the body and walk of a person and recognised the motion of an object by comparing the motion element with the model element. They recognised the motion of a person who walked straight from 30 frames of the video sequence [6]. Yaser Yacoob proposed and tested behavior modelling and a parameter model for modelling recognition. He recognised walk, march, line walk, and kick with walk, and also recognised the mean of lip change using PCA (Principal Component Analysis) linear change method from image sequence sets. His paper [7] used a natural environment and recognition areas similar to our recognition areas. However, he discussed motion recognition inside a building, and recognised the formulated motion. Also, it is presupposed that a person's movement and motion speed are similar among the motions.

2.2 Our Work

The recognition of a moving object is interesting to many researchers in digital video. This paper explores a vision-based analysis of motions that are viewed as spatio-temporal patterns. Major approaches for analysing spatio-temporal patterns include Dynamic Time Warping (DTW), Neural Nets, and Hidden Markov Models (HMMs). Neural Nets and HMMs are statistical modelling tool. DTW is a template-based dynamic programming matching technique that has been applied to problems with temporal variability. Although it has been successful in small vocabulary tasks, DTW needs a large number of templates for a range of variations. Furthermore, it cannot handle undefined patterns. HMM is a statistical modelling tool that can be applied to analysing time series with spatio-temporal variability. HMM is rich in mathematical structures; it serves as a theoretical basis for a wide range of applications. However, it has some limitations in representing non-gesture patterns. As larger data sets become available, more emphasis is being placed on neural networks for representing temporal information. Although Neural Nets have shown effectiveness in recognising static patterns (postures), they are not suited for dynamic patterns [8].

Motion recognition is based on some quantitative statistics inside a motion itself for motion classification. In this paper, we extract the features from dynamic motion objects by human knowledge and we choose the statistical weights from the features to be applied to static patterns and to handle undefined patterns. Also, we define a dynamic time interval and propose statistical weights. We propose the approach to recognition by statistical weights. Recognition by statistical weights in a dynamic time interval improves processing time with simple status analysis between neighboring frames and involves no calculation of connection between frames. We also validate our approach by implementation, testing and evaluation of the proposed algorithm.

2.3 Restrict Condition

The dynamic moving section should be indexed automatically in accordance with the motion or video scene change, but we performed manually in this paper. Detection of the player and ball, and calculation of angle and distance should have been extracted automatically, but because of the sheer volume required by this work, we performed this task manually. Automatic work is suggested for future work which must be performed.

We catch the side motion of the player. It is very difficult to exactly extract the moving object in a dynamic environment, but it is possible to do it in a limited environment and motion. We suppose that the objects are extracted in a limited environment and motion.

3.0 MOTIONS

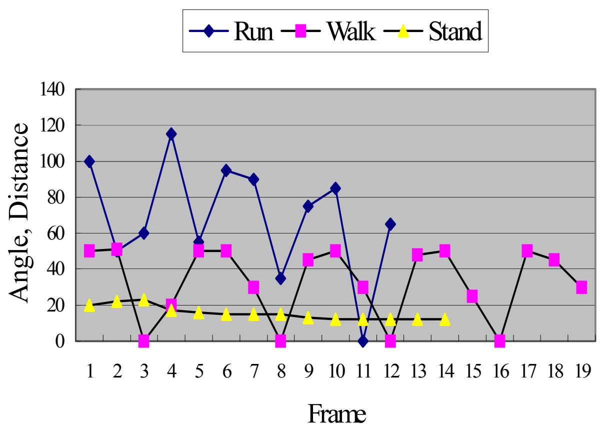
Dribble, run, walk, stand, kick, and dribble & kick are key motions in soccer games. The motion of legs can provide hints to the movement of soccer players. We can know a soccer player's motion with the activity of his lower pelvis. Fig. 2 presents a player and his leg angle.

Fig. 2: A player and his leg angle

We guess a player's motion by his pelvis, legs, and their relations. The features are extracted by his leg angle and the distance of the player to the ball. Fig. 3 is the variation graph of the angle and distance between the player and the ball. For example, Dribble: the variation graph of the angle when the player dribbles, and Dribble-b: the variation graph of the distance between the player and the ball when the player dribbles.

We choose the features to recognise the motion of the player in a video scene. The features are average leg angle, average angle difference between frames, maximum angle, distance between the player and the ball on maximum angle, distance between the player and the ball after meeting the player and the ball, and the number of times the player and the ball meet.

When we consider each feature, we see first, average angle of legs: variation of two-leg angle is a measure that can determine a player's speed. The variation such as wide legs gives a clue that a player's speed is fast, that is, the player is running. Second, the average of angle difference is considered frame by frame: when the player is not moving, the variation of two legs is very small; therefore it is a clue that the player is standing. Third, the maximum of angle: when maximum angle is big, it is a hint about how fast he is running: while the player kicks a ball, the maximum angle is wide open. Fourth, the distance between a player and a ball on the maximum angle: when a player kicks a ball, he lifts up one leg. Normally, the angle is at its maximum at this time in a motion scene. Also, when a player kicks a ball, the distance of the player and the ball must be zero.



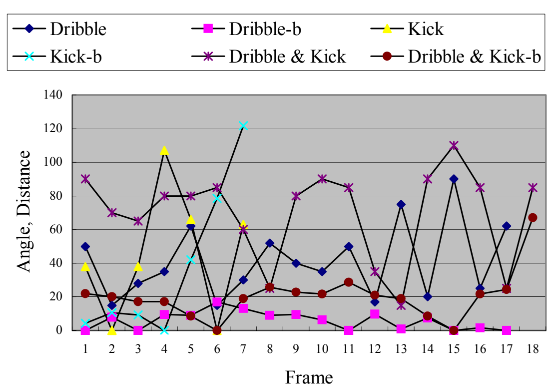


Fig. 3: Variation graphs of angle and distance between player and ball

Fifth, the distance of a player and the ball after the player and the ball meet: it is a clue as to whether the player will kick the ball or not. The action of the player's kick implies sending the ball to a great (or some) distance. Lastly, the number of times the player and the ball meet: dribbling implies that the player has a ball and is moving; therefore the player and the ball meet many times. If the distance of the player and the ball is zero just one time, it is a clue as to whether there will be a kick or a pass. Also, if the player and the ball meet many times and apart from each other, it is a clue as to whether there was a kick after a dribble. Fig. 4 expresses the features and recognition targets.

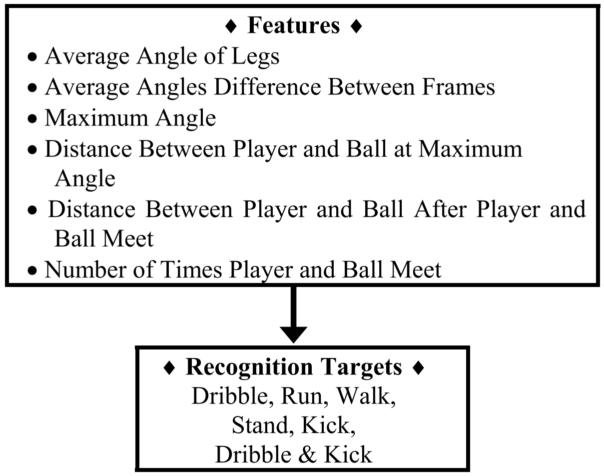


Fig. 4: Features and recognition targets

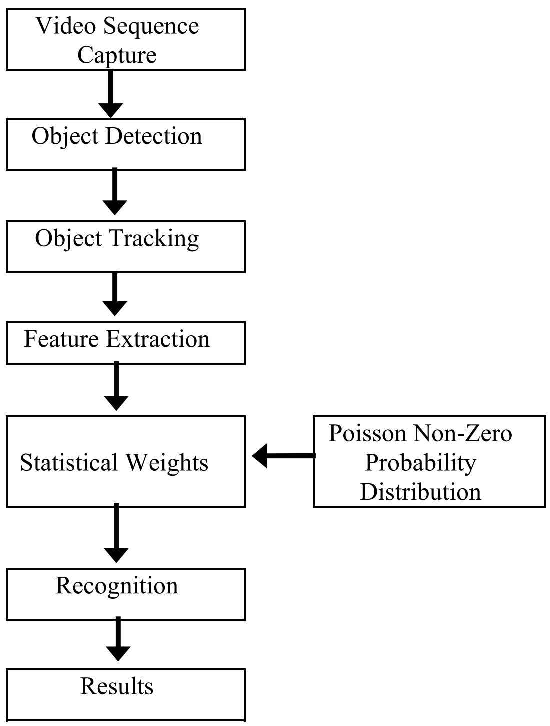


Fig. 5: Flow of moving object recognition

4.0 MOTION RECOGNITION

The start and end of a player's moving motion do not occur with a fixed time interval, but with a variable interval. Let us define a dynamic time interval. A dynamic time interval is defined as the start and end of a player's moving motion variable.

We propose an approach to motion recognition of the player by statistical weights. Fig. 5 is a flow diagram of motion recognition. We use six features defined in Fig. 3 for motion recognition of the player.

4.1 Probability Distribution

4.1.1 Poisson Distribution

Poisson distribution states that an event number occurs in a given time or space. Poisson distribution is satisfied following two conditions. First, the advent of the event is the same in a given time or space. Second, in a given time, the advent of the event is independent of the advent of the event in other given times. This distribution is expressed by the following function:

$f(x) = (\mu e^{-\mu})/x$£¡
$(x = 0, 1, 2, 3, \dots)$

x is the probability variable, μ is the number of events, and e = 2.71828

4.2 Motion Recognition of the Player

4.2.1 Statistical Weights

We propose an approach to recognition by statistical weights on non-zero Poisson probability distribution. Recognition by statistical weights is that which endows the weights to the features. Also, we recognise the target result in accordance with the weights multiplied by each recognition target. The largest value is our recognition target. Fig. 6 illustrates the statistical weights recognition concept.

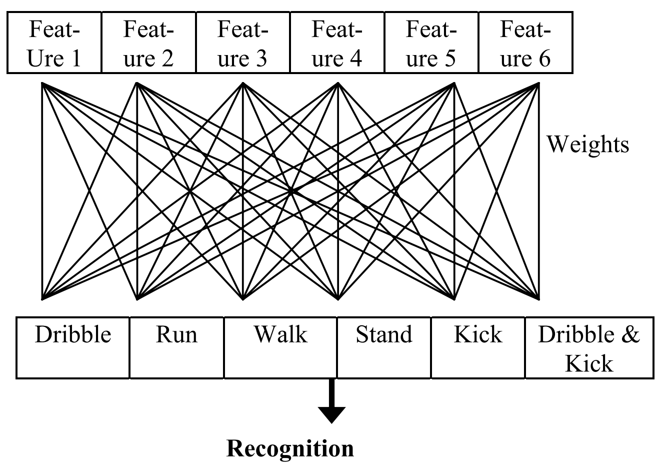


Fig. 6: The concept of recognition by statistical weights

Let us consider our approach. We divide the features by sections, and we express a section as following:

$(y-1) * 10 <= y$ section $< y * 10$
(but, $1 <= y <= 18$ and y is an integer)

However, the number of times the player and the ball meet is divided by
$y-1 <= y$ section $< y$

For example, the legs' angle is 46.7 on the run, therefore it is distributed using Poisson fifth section. In accordance with the distribution, if the probability of zero is A in the fifth section, then the probability of greater than zero is 1.0 – A in the fifth section. It is a weight. For example, if the Poisson distribution value of zero for the average angle of legs of running between 40.0 and 49.9 is 0.2, then the probability of greater than zero is 0.8, that is, the probability weight is 0.8. If μ is average in y section, then the Poisson probability of zero expresses such as the following function:

$f(x,y,\mu) = (\mu^x e^{-\mu})/x!$
$(x = 0)$, (μis average in y section),

So, we can represent statistical weights by non-zero Poisson probability distribution to the next function.

$$W(y) = 1.0 - f(x, y, \mu)$$

On the other hand, if we define Min(f(x, y, μ)) as the minimum of f(x, y, μ), (y>1), we present the distance between the player and the ball after the player and the ball meet, and the number of times the player and the ball meet in the following function:

$$W(y) = 1.0 - \mathrm{Min}(f(x, y, \mu)), (y>1)$$

Table 1 is the statistical weights table. If W(I, k) is a feature respectively in a human motion (but I, k=1,2, … 6), and targets are Dribble, Run, Walk, Stand, Kick, and Dribble & Kick (e.g., Target 1 : Dribble, Target 2 : Run, Target 3 : Walk, Target 4 : Stand, Target 5 : Kick, Target 6 : Dribble & Kick). We present the result as the following function:

$$\text{Result} = \text{Maxtarget}(i, j)\,(W(I*1, j) * W(I*2, j) * \ldots * W(I*6, j)\,)_{i=1}^{i=18} \quad (i=1, 2, \ldots, 6)$$

Table 1: Statistical weights table

Section **Target**	**1**	**2**		**18**
Target(1) … Target(6)	W(1,1) … W(6,1)	W(1,2) … W(6,2)	…	W(1,18) … W(6,18)
Target(1) … Target(6)	W(7,1) … W(12,1)	W(7,2) … W(12,2)	…	W(7,18) … W(12,18)
Target(1) … Target(6)	W(13,1) … W(18,1)	W(13,2) … W(18,2)	…	W(13,18) … W(18,18)
Target(1) … Target(6)	W(19,1) … W(24,1)	W(19,2) … W(24,2)	…	W(19,18) … W(24,18)
Target(1) … Target(6)	W(25,1) … W(30,1)	W(25,2) … W(30,2)	…	W(25,18) … W(30,18)
Target(1) … Target(6)	W(31,1) … W(36,1)	W(31,2) … W(36,2)	…	W(31,18) … W(36,18)

5.0 EXPERIMENTAL RESULTS

5.1 Sampling of Data

Data in this paper are sampled from the final game and decisive games of the third and fourth winners in the 1998 France World Cup. The final game is France versus Brazil, and the decisive game is Netherlands versus Croatia. Also, we sampled from an international goodwill match of Korea versus Brazil in Seoul, 1999. These data are recorded by videotapes. We extracted 5,650 frames from the videotapes. These are 360 scenes to recognition, 60 scenes for each motion.

5.2 Experimental Results

We use 20 training scenes on each motion, a total of 120 scenes as the statistical data for statistical weights. We use 40 testing scenes on each motion, a total of 240 scenes to test.

The following results are examined. For the recognition by statistical weights in Table 2, two dribbles are non-recognised as anything. Six runs are mis-recognised as walks. Walk and stand are perfectly recognised. One kick is recognised as a dribble & kick and eleven kicks are non-recognised as anything. Also, ten dribble & kicks are recognised as dribbles.

Table 2: Recognition by statistical weights

	Mis-Recognition	**Non-Recognition**	**Recognition Rate**
Dribble	0	2	95%
Run	6 -> Walk	1	82.5%
Walk	0	0	100%
Stand	0	0	100%
Kick	1-> Dribble & Kick	11	70%
Dribble & Kick	0	10	75%
Total	7	24	87.1%

5.3 Evaluation

We test the proposed approach to recognition by statistical weights. We obtained the recognition result of 87.1%. The numbers for run show many mis-recognitions as walks. The reason why run is recognised as walk is because running features are not expressed clearly when the run is very slow, because the legs' angle is not wide open which is similar to fast walking. The reason why walk is recognised as stand is because a walk can be very slow, so walk is confused with standing. Walk and stand are recognised correctly. The walk and stand features are clearly distinguished from others. Dribble, kick, dribble & kick have much non-recognition. The reason of mis-recognition is that dribble, kick, and dribble & kick are not clearly distinct in relation to the player and the ball.

6.0 CONCLUSION

This paper introduces the approach to the player's motion in a dynamic video scene. Our work has been tried and tested on soccer games. Dribble, run, walk, stand, kick, and dribble & kick are major motions in soccer games. Time interval in a dynamic video scene is defined and features chosen. We proposed statistical weights to recognise the motions of moving objects in a video scene. Features by human knowledge are chosen, such as average leg angle, average angle difference between frames, maximum angle, distance between the player and the ball on maximum angle, distance between the player and the ball after the player and the ball meet finally, and the number of times the player and the ball meet. We extracted static patterns from dynamic motion objects in a recorded video scene of a soccer game, and we described the extracted static patterns to statistical weights. Recognition by statistical weights in a dynamic time interval improves processing time by simple status analysis between neighboring frames and requires no calculation of connection between frames.

We have proposed the statistical weights as an approach to recognition. Tested result was 87.1% successful by the proposed approach. It was tested on soccer games in a recorded video scene. Our future works will be on automatic indexing for the dynamic section of a moving object, automatic detection of objects and extraction of features, and the approach to object recognition by Neural Nets.

REFERENCE

[1] Elisabeth Andre, Gerd Herzog, and Thomas Rist, "On the Simultaneous Interpretation of Real World Image Sequences and Their Natural Language Description: The System SOCCER", in *Proc. of the 8th ECAI, Munich*, 1988, pp. 449-454.

[2] Chueh-Wei Chang and Suh-Yin Lee, "A Video Information System for Sport Motion Analysis". *Journal of Visual Languages and Computing*, 1998, pp. 265-287.

[3] Dennis Yow, Boon-lock Yeo, Minerva Yeung and Bede Liu, "Analysis and Presentation of Soccer Highlights from Digital Video", in *ACCV '95*, 1995.

[4] Koh Kakusho, Noboru Babaguchi and Tadahiro Kitahashi, "Recognition of Social Dancing from Auditory and Visual Information", in *International Conference on Automatic Face and Gesture Recognition, IEEE*, 1996, October 14-16.

[5] Lee Campbell and Aaron Bobick, "Using Phase Space Constraints to Represent Human Body Motion", in *International Workshop on Automatic Face and Gesture Recognition, Zurich*, 1995.

[6] Jia-Ching Cheng, and Jos M. F. Moura, "Model-Based Recognition of Human Walking in Dynamic Scenes", in *IEEE Signal Processing Society 1997 Workshop on Multimedia Signal Processing, Princeton, New Jersey, USA, Electronic Proceedings*, June 23-25, 1997.

[7] Yaser Yacoob, "Parameterized Modeling and Recognition of Activities". *Computer Vision and Image Understanding*, Vol. 73, No. 2, February 1999, pp. 232-247.

[8] Hyeon-Kyu Lee and Jin H. Kim, "An HMM-Based Threshold Model Approach for Gesture Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 10, October 1999.

BIOGRAPHY

Nam Sung Hyun is a PhD candidate at the AI lab of the University of Korea. His research interests include image processing, AI, multimedia. He obtained his master's degree in computer science from the Hankuk University of Foreign Studies in 1991. He is a member of Korea Information Processing Society.

Kim Hyun Sook is a professor of computer science from Sinsung Technical College. Her research areas include image processing and computer graphics. She obtained her Master of Computer Science from Sugang University in Korea. She is a member of KISS (Korean Information Science Society).

Hwang Chong Sun obtained his PhD of Statistics and Computer Science from University of Georgia in 1978. He is a professor of computer science from Korea University. His research interests include computer algorithm, database, and distributed system. He is a member of KISS.

Yang Young Kyu obtained his PhD of image processing from University of Texas A&M in 1984. He is currently working in the ETRI (Electronics Telecommunications Research Institute). His current research interests include image processing, pattern recognition, and remote sensing. He is a member of the Korean Information Science Society.