Prediction of Basketball Free Throw Shooting by OpenPose

Masato Nakai,Yoshihiko Tsunoda, Hisashi Hayashi, Hideki Murakoshi

School of Industrial Technology, Advanced Institute of Industrial Technology
1-10-40, Higashiooi, Shinagawa-Ku,Tokyo, 140-0011, Japan

Abstract. OpenPose, which was developed by Carnegie Mellon University (CMU) presented at CVPR 2017 last summer, takes in real-time motion images via a simple web camera and is capable of recognizing skeletons of multiple persons in these images. It also generates recognized skeleton point coordinates to files. OpenPose is featured by CMU's original top-down method for real-time recognition and it is open online especially for research purposes. Thus we aimed to build a posture analysis model using OpenPose skeletal recognition data and verifying the practicality of OpenPose by verifying the accuracy of the model. As a posture analysis model, we adopted a logistic regression model that predicts the shooting probability of the basketball free throw with skeleton posture data as explanatory variables and the shooting or not as a target variable. As the result, high prediction accuracy was obtained. Therefore, posture analysis using OpenPose has been verified to be practical with our model. We consider that with many skeleton data which are easily provided by a simple web camera, OpenPose make statistical diagnostic approach possible. We also consider it could lower costs (in both financial and time-wise) of such analysis which has previously required more equipments and more time for preparation regarding motion capture analysis systems.

Keywords: OpenPose, Logistic regression, Basketball shooting prediction, Posture diagnosis

1 Introduction

We consider that a large amount and variety of human posture data with high precision are required to improve performance in statistical posture analysis. However we found that there are little available posture data because of complexity for acquiring posture data. In order to acquire the whole posture data as a time-series, there are major two methods. One is video motion analysis(manual marking) and the other is motion capture analysis. Video motion analysis, which divides into photographs and marks annotations on the pose images, is very laborious. Motion capture analysis is necessary to attach so many sensors on human limbs that acquisition of data is complicated in operation. At CVPR 2017 conference, Carnegie Mellon University (CMU) presented OpenPose[1][1] which can recognize skeletons of multiple players in real-time, using a simple web camera, as shown in Fig. 1. OpenPose adopted the unique top-down position recognition using Deap Learning and also the unique algorithm as affiliation recognition of body parts by PAF (Part Affinity Fields)[2]. As a result, in the moving skeletal pictures generated by OpenPose, the skeleton marks are shown and overlapped well with the figure of people, and it seems that recognition accuracy is very high even for various people in various environments. OpenPose can be expected as a convenient generation model of available and accurate posture data. In oder to evaluate performance of OpenPose, we built a basketball shooting prediction model as a prototype using real-time skeletal data generated by OpenPose. As the result, we found that the shooting prediction model showed sufficient high accuracy. Thus, we found that OpenPose is a convenient and practical generator of posture data. The rest of this paper is structured as follows. In Section 2, we briefly review of previous posture analysis method. In Section 3, we define the experimental methods. In Section 4, we show experimental results. In Section 5,we conclude this paper. In Section 6, we present future work.

Fig. 1. OpenPose

2 Previous research for posture analysis

As a previous analysis of sports motion, there was a method called video motion analysis which divides video into photographs and marks points on the pose image for annotation[3]. As a direct sampling posture data, motion capture was used to collect data from sensors on body and limbs[4]. However, these methods are so expensive to collect data that statistical models such as regression could not be applied. Thus, these method were often used for argument about the difference of posture between beginners and experts. On the other hand MicroSoft

[1] OpenPose realizes three-dimensional acquisition by stereo (compound eye) camera in March 2017, but in this research, OpenPose of 2D position recognition version using monocular Web camera is used because of easy operation and sufficient use frequency.

KINECT, which is not sold now, can easily take 3D posture data, but the sensing range is very narrow and recognition accuracy of skeletal point is low[5].

As a statistical approach, it is realized that winning prediction of basketball game was modeled by logistic regression using records which include the winning / losing result and the frequency of shooting and robbing ball in the game[6].

As a time-series analysis for motions, there was a research that tried to transfer abstracted motions from a human to a robot with hidden variables estimated by Hidden Markov and reversely predict the next action of the robot from estimated hidden variables [7]. However our shooting prediction of basketball free throw is not a general time-series model that predicts the next action from the last time-series of motions because our model predicts a result whether to shoot in the basket or not rather than an aciton. The number of persons monitored in this experiment was limited to 51. We adopted a logistic regression[8] using the features which are composed with the position of start and end or the difference, velocity and acceleration between start and end position extracted from a time-series of the free throw motions.

3 Method

3.1 Subject of Experiment

In order to construct the shooting prediction model of basketball free throw, we took movies about basketball free throw motion with web camera. The subject of Experiment covered various skill levels of 51 people. We generated their skeletal data by OpenPose from the movies. In the experiment, twenty of 51 people succeeded in free throw.

3.2 Output of OpenPose

The version of OpenPose[1] adopted in this paper is for 2 dimensional skeleton recognition, and the skeletal coordinates of 18 points shown in Fig. 2 are outputted to files by about 10 to 20 frames per second depending on a computer performance[2] and connected as shown in Fig. 3 to make time-series data. A skeletal coordinate is composed of 3 values which are x as horizontal, y as vertical and p as confidence probability. We ignored low confident coordinates with less than 0.7 confidence probability.

[2] Our experimental machine is CPU: AMD Ryzen 7 1800X, MEMORY: 16GB, GPU: NVidia GeForce GTX 1080ti,OS: Ubuntu 14.04 LTS,CUDA version: 8.0, cuDNN version: 5.1 for CUDA8.0

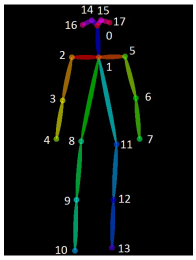

Fig. 2. Skeleton points of OpenPose referred from [2]

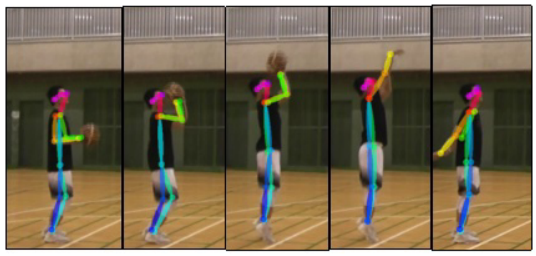

Fig. 3. Connection of frames

Though we used only skeletal data for prediction of shooting model, OpenPose can recognize also hands and faces as shown in Fig. 4 and outputs each recognized data to each file.

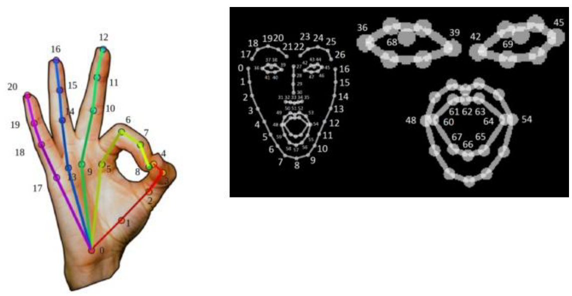

Fig. 4. Recognized finger and face points of OpenPose (cited from [2])

3.3 Statistic model

The shooting prediction model is a binary prediction as to whether to enter a basket or not. As major binary prediction models there are logistic regression[8], SVM[9] and Xgboost[10]. The SVM using the kernel method is a nonlinear model which may make high accuracy but cannot calculate the shooting probability. Xgboost using the stochastic gradient method that has a high reputation for accuracy and robustness is not adequate for a diagnostic model because this model cannot indicate explicitly the degree of importance of explanatory variables. So we adopted a logistic regression model as shown in Fig. 5 that is one of most used in data analysis.

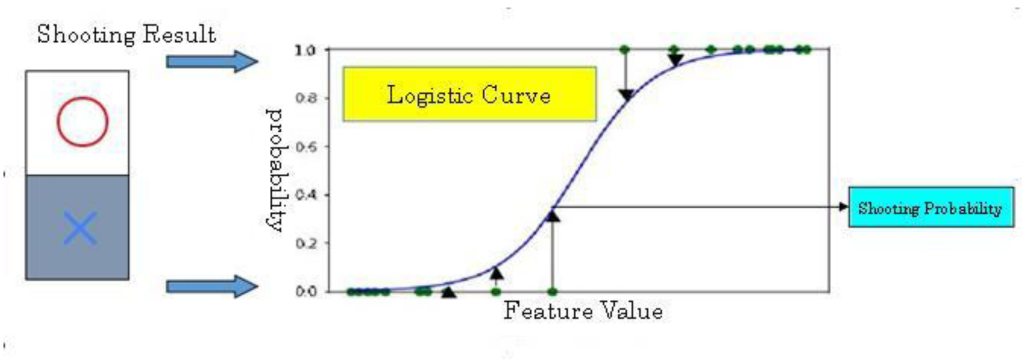

Fig. 5. Logistic regression

Probability of logistic regression is as follows using multiple regression coefficients β and features.

$$\mathcal{Z} = \sum_{i=1}^{n} \beta_i \cdot feature_i \tag{1}$$

$$probability = \frac{1}{1 + \exp(-\mathcal{Z})} \tag{2}$$

3.4 Valuable of Logistic regression

The explanatory variables of the logistic regression model were features calculated from the time-series data. Features include the positions of skeletal points, moving speed, acceleration, etc. The target variable is the fact whether a shooting ball entered into the basket or not. In the time-series data, as shown in Fig. 6, everyone bent knees at the start of the throw and lifted hands to the highest level upon completion of the throw. So we decided to define the interval of a free throw between the time point when the knees were bent the most as its start and the time point when the hands were lifted to the highest physical point as its end. All the coordinate positions were relative from the neck point.

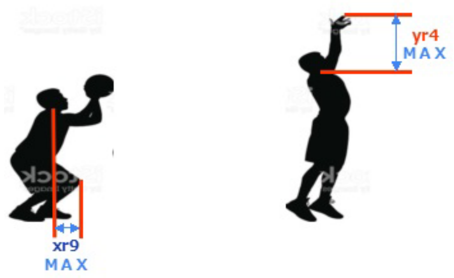

Fig. 6. Posture of start and end

4 Result

4.1 Accuracy of Logistic Regression

In general, the precision of the logistic regression model is expressed in the pareto diagram in Fig. 7. In this diagram, the horizontal axis shows the composition rate of all the members in descending order of shooting probability, and the vertical axis shows the composition rate of the number of people who succeeded in shooting into the basket. Red dots in Fig. 7 indicate number of accumulated people who shot in basket. About 40% people succeeded in shooting into the basket at this experiment. If the model was perfect, it would be represented by the line of the perfect model with descending order of shooting probability, and if the shooting probability of the model was uncorrelated with the actual shooting in basket, it would be the line of the uncorrelated model. The accuracy of logistic regression model is indicated by the ratio of the area A of the cumulative curve shown in Fig. 8. This figure shows that high accuracy AR = 41% was obtained.

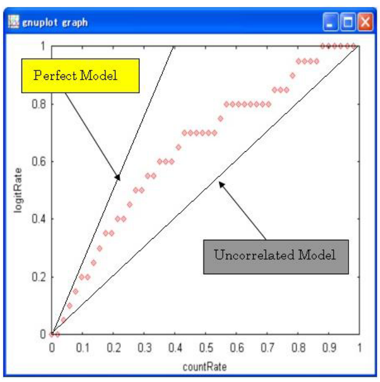

Fig. 7. Parete figure

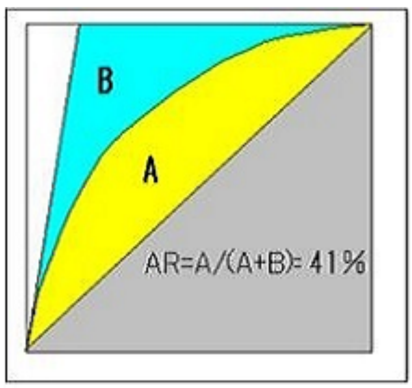

Fig. 8. Area of AR value

4.2 Interpretation of significant features in Logistic Regression

As a result, significant features that make the high shooting probability in the logistic regression were shown in Fig. 9. The shooting probability becomes higher when the blue color features are larger. The shooting probability also becomes higher when the red color features get smaller. From this result, the following was found out. It shows that the shooting probability is higher if the bend of the knee is increased and knees are pulled quickly and at the same time the ball is pulled and thrown over head. This motion uses the force of the knee extension and the centrifugal force created when throwing the ball overhead.

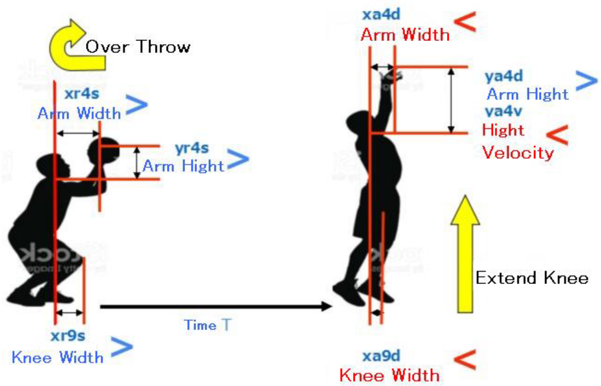

Fig. 9. Relation for shooting probability and features

4.3 Real time Diagnosis

The diagnostic system using OpenPose can display the shooting probability in real-time as shown in Fig. 10. and even if there are not any basket and ball, it becomes possible to judge the skill level directly just by gesture.

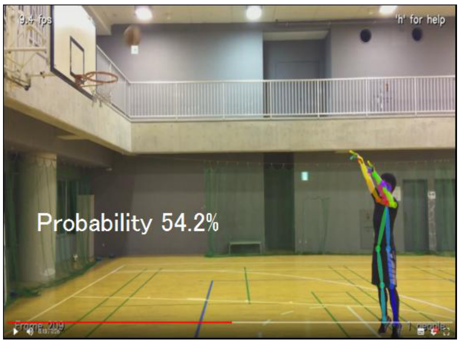

Fig. 10. Real time diagnosis by OpenPose

4.4 Posture diagnosis

By comparing feature quantities between beginners (people with low shooting probability) and experts (people with high shooting probability), it is possible to diagnose the amount of correction for beginner postures. In the example of Fig. 11, one of the remarkable differences between beginners and experts is the position of the arm at the start. The beginners pushed the ball from the chest but the experts put the ball in front of the head and threw the ball over head. In this case, it is necessary to teach the beginners the form of overhead throwing.

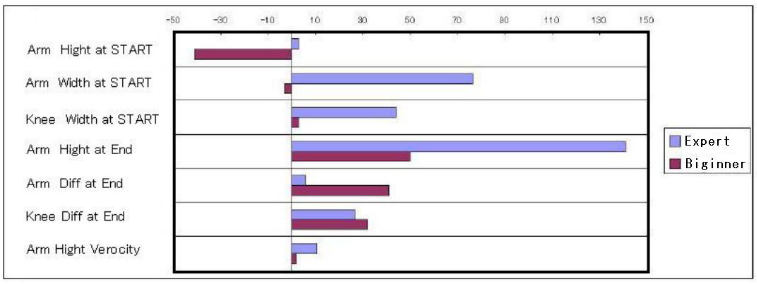

Fig. 11. Comparison of features between beginner and senior

5 Conclusion

As a result, high accuracy was obtained in logistic regression model, the following was found out.

1. The skeletal data recognized by OpenPose were found to be highly applicable with sufficient accuracy.

2. In the previous posture diagnosis, data were generated by marking on a picture frame from a video stream or collected sensor signals by motion capture on the human parts. These methods were so expensive for data collection that statistical models could not be introduced. On the other hand, OpenPose can easily collect data by using a simple web camera, it became possible to obtain more accurate posture diagnosis by collecting more data.

6 Future work

The data of basketball free throw in this experiment were taken from one side only by a web camera, so it was suitable to analyze with 2 dimensional data provided by OpenPose. However analysis of general sports motion requires 3 dimensional data like a tennis or ballet dance, so it is necessary to use 3 dimensional OpenPose or expand 2D data generated by 2D OpenPose to 3D data[11].

Acknowledgment

We would like to thank Basketball Club Team of Tokyo Metropolitan College of Technology for taking movie of basketball free throw. We would like to special thank to Dr. Atushi Shibata of AIIT for provision of experimental computation environment.

References

1. CMU-Perceptual-Computing-Lab,https://github.com/CMU-Perceptual-Computing-Lab/openpose
2. Cao Z., Simon T., Wei S., Sheikh Y. : Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields,CVPR2017
3. Hulka K., Cuberek R., Svoboda Z . : Timemotion analysis of basketball players: a reliability assessment of Video Manual Motion Tracker 1.0 software,Journal of Sports Sciences, 2013
4. Wang X.. : An Optimization Method of Basketball Teaching and Training System Design based on Motion Capture Technology,2017
5. Yokota K. : Construction of Motion Analysis System using Kinect
6. Shanahan J. K. : A model for predicting the probability of a win in basketball,University of Iowa 1984
7. Inamura T., Nakamura Y., Toshima I.. : Embodied symbol emergence based on mimesis theory, International Journal of Robotics Research,vol. 23, no. 4, pp. 363377, 2004.
8. Bishop C. M. : Pattern Recognition and Machine Learning 4.Linear Models for Classification,Springer(2006)
9. Platt J. C. : Fast Training of Support Vector Machines using Sequential Minimal Optimization,1999
10. Chen T., Guestrin C., Christopher : XGBoost: A Scalable Tree Boosting System,Arixv2016

11. Kudo Y., Ogaki K., Matusi Y., Odagiri Y. : Unsupervised Adversarial Learning of 3D Human Pose from 2D Joint Locations,arXiv:1803.08244
12. Kempka M., Wydmuch M., Runc G., Jaskowski W. : ViZDoom: A Doom-based AI Research Platform for Visual Reinforcement Learning , 2016 IEEE Conference