Sports Skill Frequency Analysis with Motion Image Data

Toshiyuki MAEDA
Department of Management Information
Hannan University, Japan
Email: maechan@hannan-u.ac.jp

Masumi YAJIMA
Department of Economics
Meikai University, Japan
Email: myajima@meikai.ac.jp

Akiyoshi WAKATANI
Department of Intelligence and Informatics
Konan University, Japan
Email: wakatani@konan-u.ac.jp

Abstract—This paper addresses skill frequency analysis with motion image data of volleyball attack. We try to justify a assumption that expert skills have relatively low frequency motions rather than novice skills as the similarity of human postural control. For this purpose we have experiments and analyze sports skills as for frequency of motion using time series motion images of volleyball attacks. In our research, volleyball play is analyzed with motion image data recorded by hi-speed cam-coder, where we do not use physical information such as body skeleton model. Time series data are obtained from the motion image data with four marking points, and analyzed in terms of motion frequency. As the experiment results, we in some content classify expert, middle, and novice skill image using frequency data.

Keyword: Time series data; Sports skill; Motion image; Frequency analysis

I. Introduction

For engineering skills as well as sports skills, most of researches treat structure models retrieved from body and skeleton information, for example, activity or bio-mechatronical data [7]. Those are possibly because of they believe that technical skill models consist of some skill hierarchies, which are environmental adjustment, human intention, and so on
[10].

A research of Matsumoto et al. presents crafts-mens skill architecture model, and in the paper crafts-men select action procedure adapted with environments [6]. It is as though hard for expert workers to clarify their models by themselves. They act unconsciously and achieve expert skill technique with their own models.

On the other hand, there are some researches for human postures. For instance, Nomura et al. [9] present that reaction force of human stance on the ground is modulated in hemodynamic cardiac cycle synchronously. They also imply that deterministic slow oscillations, which have intermittent features of control strategy, might be a basic mechanism to generate postural sways.

This suggests similar postural control may exist in sports skills as well as usual motion control, and that means expert skills have lower frequency motion comparing to novice one. We had researched sports skills, and classify skill models by motion image data analysis. This paper addresses personal sports skill classification focused on volleyball play motion frequency.

II. Background

As for the background of this research, we introduce some related works, and our previous works.

A. Related Works

A research by Kirschenbaum et al. [2] is that positive self-monitoring, more than the negative monitoring or control processes, might improve average scores of bowling for sixty unskilled league bowlers, by self-regulated laboratory researches.

However, negative self-monitoring seemed to produce relatively the best result for sixty-seven skillful league bowlers. In partial support of the assumptions, positive self-monitoring makes significant progress for bowling averages, more than all other groups of low-skilled bowlers. In addition to other results in cognitive therapy and sports psychology, these results suggest they describe the circumstances in which the selfmonitoring facilitates positively with self-regulation.

Smeetona et al. [11] show that the performance in team sports is based on the ability to recognize patterns of plays. Skills of pattern recognition might reuse from a sport to another one when there are some similarities in strategies, relations, and perceptual features. They might encode and decode useful information. In this research, expert and less expert volleyball, field hockey and soccer players' abilities for pattern recognition are investigated.

Subjects watch well-formed and ill-formed play sequences for all sports, and half of them are randomly showed with clips not previously viewed. Here they have to identify previous action sequence quickly and accurately. Pattern recognition skills are dependent to transfer on the skills of participants, practiced sports, the essence of tasks and extent of structures. In this research, expert soccer and hockey players are quicker than volleyball players to recognize soccer and hockey action plays. Performance is not different from the well-formed volleyball trials between the expert volleyball, field hockey and soccer players to each other. The expert field hockey and soccer players can reuse perceptual information or strategies among the respective sports. The results of novice participants are not so relevant. They discuss the transfer and diversity cross-over the domains for domain-specific skills.

Nomura et al. [9] report human being modulates ground reaction force with the quiet stance with cardiac cycles

978-1-5386-2615-3/18/$31.00 ©2018 IEEE

synchronously. In this research, periodic hemodynamic force almost causes small disturbance torque to ankle joints, and they are thought as a source of internal perturbation which causes postural sway. They hence consider dynamics of postural sway for reversed pendulum models for intermittent control strategy by comparison to usual feedback controller in continuous time sequence. They have investigated that each control model can exhibit human-like postural sway, featured by the behavior at low frequency band, on weak perturbation by periodic and randomly mimicking forcing hemodynamics.

They suggest that not continuous control with typical feedback gain but intermittent control exists in the human-like sway pattern. Furthermore, for generating the postural sway, there might a certain mechanism for deterministic, including chaotic, slow oscillations, and those are by intermittent control strategy in addition to the tiny hemodynamic perturbation.

Those researches are the basis of our researches, and among them, Nomura's work especially inspires us much and then this paper focuses on motion frequency.

B. Our Previous Works

We have researched some of sports skill analyses. In [4], we focus on table tennis action, especially fore-hand strokes, as sports play, and classify skill models by time series motion image data analysis without physical structure models. We had investigated three play levels as expert, intermediate and novice, and try to discriminate the models by knowledge retrieval technologies. This research is a attempt to retrieve skill models only from motion image data. In this paper, we carry out experiments and the results suggests that the expert or intermediate subjects may group some categories as for technical skills, but there might not be "novice category," and that is because novice players' skills varies to each other. Moreover, reconstruct time series data are reconstructed from the original data, and apply the data into knowledge retrieval techniques as J48 (a C4.5 implementation), Naive Bayes Tree, and Random Forest. The recognition rate to evaluate data is quite bad, but that of J48 is better as for learning and test than the others. We thus make an attempt for two-class analysis, and that implies Intermediate class might be sorted as Novice class.

In [5], we present sports skill discrimination by motion image data focused on volleyball attack skill. This research attempts to certify the assumption that expert skills have relatively low frequency motions rather than novice skills as the similarity of human postural control. For this purpose we have experiments and analyze sports skills as for frequency of motion by time series motion pictures of volleyball attacks. In this paper, volleyball play is analyzed with motion image data recorded by hi-speed cam-coder, where we do not use physical information such as body skeleton model, and so on. Time series data are obtained from the motion image data with four marking points, and analyzed by Fast Fourier Transform (FFT) and clustering data mining method. As the experiment results, we have found that y-axes of the novice data have much more high-frequency data, and that implies novice motions might have high frequency motions, and that may support our assumption.

Based on these researches, this paper furthermore addresses analysis on motion frequency in various methodologies.

III. Mel-frequency cepstral coefficients

For analysis of motion frequency, we introduce Melfrequency cepstral coefficients (MFCCs) [3], [8] in addition to Fast Fourier transform.

The mel-frequency cepstrum (MFC) is representations of the short-term sound power spectrum, and it is based on a log power spectrum with linear cosine transform, and thus on nonlinear mel scale of frequency in the field of sound processing.

MFCCs are coefficients which make up an MFC collectively. Those are derived from type of cepstral representations of audio clips, or a nonlinear "spectrum-of-a-spectrum."

The cepstrum and mel-frequency cepstrum is the most significantly different from the frequency bands in the MFC, and they are equally spaced on the mel scale. That looks after the response of human auditory system more closely than the linearly-spaced frequency bands in the normal cepstrum. For example, this frequency warping may show for better sound representation in the field of audio compression.

The Mel scale relates pitch, or perceived frequency, of pure tones to the actual measured frequency [1]. Human being is much better to discern small changes in pitch at low frequencies rather than at high frequencies. To introduce this scale may make the features fit more closely what human being listens.

The formula to convert from frequency to Mel scale is as follows:

$$M(f) = 1125\ln\left(1+\frac{f}{700}\right)$$

To get from Mel scale back to frequency is as follows:

$$M^{-1}(m) = 700\left(e^{\frac{m}{1125}} - 1\right)$$

Brief steps to calculate MFCCs are as follows:

1) Make frames of the signals into short ones.
2) Take the periodogram estimate of power spectra for all frames.
3) Apply the mel filterbank to each power spectrum, then sum the energy for all filters.
4) Calculate logarithms of all filterbank energies.
5) Calculate DCT of all log filterbank energies.
6) Keep DCT coefficients two to thirteen, then discard the rest.

We introduce this methodology in addition to FFT as variation of feature retrieval from motion frequency data.

IV. Experiment and Discussion

A. Experiment condition

We here make a target on volleyball attacks, same as the previous research [5], and make further analysis of volleyball attack. For this analysis, we use the same subjects who are six university students. In terms of skill analysis as representation of plays, we set up three levels as below;

- Expert class: members of university volleyball club,
- Middle class: ex-members of high or junior high school volleyball club, and
- Novice class: inexperienced students.

Each class has two subjects players on this experiment. All subject players are marked and measured at four points as:

- Right elbow
- Right shoulder
- Right waist • Left knee

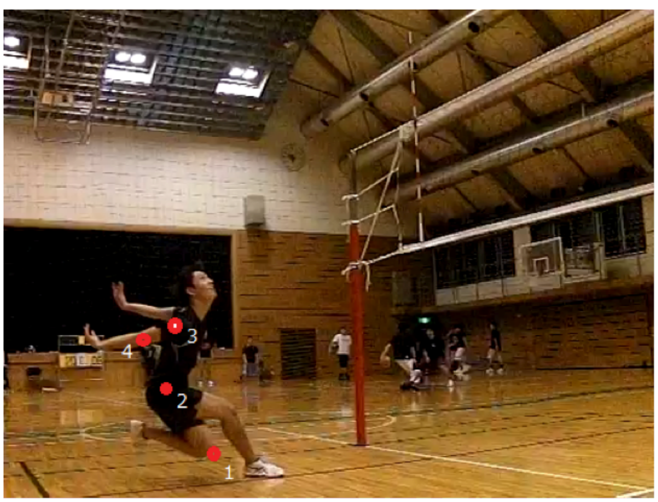

FIGURE 1. Measurement markings.

Figure 1 presents the positions of markings for measuring. We record several motion images, or movies, for swing traces of attacks by a cam-corder. The recording situation is as follows:

- The cam-corder has a resolution of 512 × 384 pixels, and a frame-rate of 300 frames per seconds.
- It is installed besides of the players.
- Subjects play in several minutes, and during that time, some attack motions are recorded for each subject player.

B. FFT and MFCCs Analysis

In this research, we have an attempt to classify six subjects to each other. 300 frames (one minute) in the recorded movies are extracted from the start of the take-back till the end of the attack motion as well as previous research. After that, We retrieve two dimensional axes position data of four marking points for each frame shown as pixel values. The beginning point is fixed, for the basis of the coordinates, at the shoulder position of the first frame. We reconstruct feature vector data from time series data of each axis by MFCCs as well as FFT.

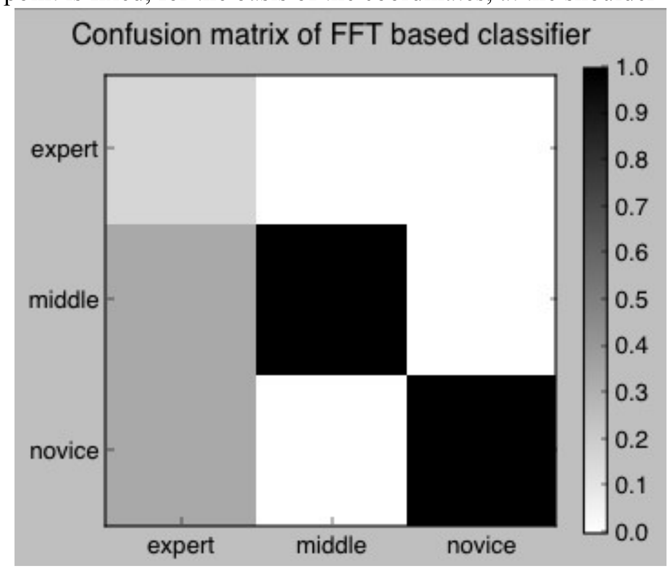

FIGURE 2. Confusion matrix of FFT based classifier.

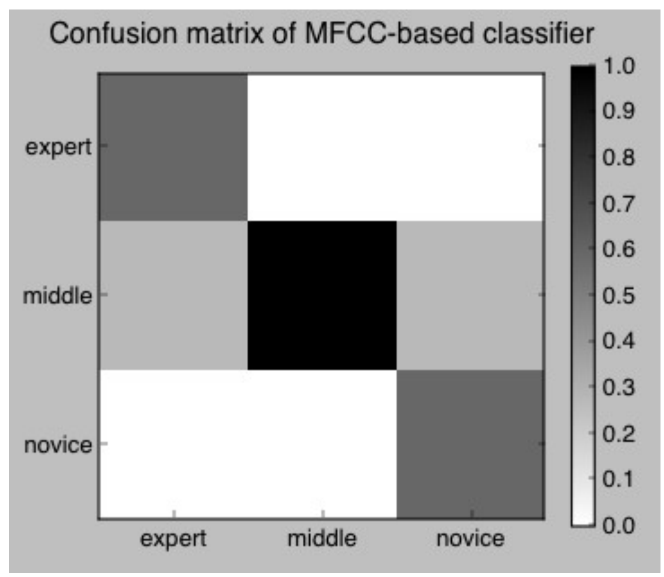

FIGURE 3. Confusion matrix of MFCC-based classifier.

In the previous research [5], we use hierarchical clustering, but in this research, we attempt to classify each subject by logistic regression with cross validation.

The reason to apply logistic regression to this analysis is because:

- Clustering is a task to group sets of objects by similarity to each other.
- Logistic regression is one of appropriate regression analyses for conduction of dependent variables.
- One of objectives of this research is to investigate the possibility of classification for multi dependent variables data, and then logistic regression should be fit better than clustering.

Figure 2 is the classification result by the FFT-based data. This diagram indicates the result is not bad as for classification, especially 'middle' and 'novice' is classified correctly, though 'expert' class is not so good. The accuracy rate is 73%. Figure 3 is the classification result by the MFCC-based data. Comparing to the above FFT-based classification, this result does not show so good, but a little better than the above result by FFT, especially as the accuracy rate is 77%.

In both cases, 'middle' is classified more clearly rather than the others, and that may indicate the skill of 'middle' is quite special, though further experiments are needed as we don't think we have enough amount of data for this sort of analysis.

V. CONCLUSION

This paper presents sports skill classification by motion image data for volleyball attacks. Volleyball play is analyzed with motion image data recorded and time series data are obtained from motion image with marking points. We analyzed using MFCCs and FFT methods, and then logistic regression techniques, and we have the better classification result by MFCCs than that by FFT, and that suggests a possibility of skill research methodology by MFCCs, though further investigation is required.

As for the future plans, we should carry out more experiments and investigate more precise data for classification. Moreover, we should attempt to use other classification technologies such as Support Vector Machine (SVM), Self Organization Map (SOM), and so on.

ACKNOWLEDGMENT

Part of this research was supported by JSPS KAKENHI Grant Number 15K02185 and 15K03802. This research was partially assisted, especially for data management, by Mr. Y. Tamari and Mr. Y. Tsujino at Hannan University. The authors greatly appreciate those.

REFERENCES

[1] Practical Cryptography. Mel frequency cepstral coefficient (mfcc) tutorial. http://practicalcryptography.com/miscellaneous/machinelearning/guide-mel-frequency-cepstral-coefficients-mfccs/, 2016.

[2] A. M. Ordman D. S. Kirschenbaum, A. J. Tomarken, and R. Holtzbauer. Effects of differential self-monitoring and level of mastery on sports performance: Brain power bowling. *Cognitive Therapy and Research*, 6(3):335–341, 1982.

[3] B. Logan et al. Mel frequency cepstral coefficients for music modeling. In *ISMIR*, 2000.

[4] T. Maeda, M. Fujii, I. Hayashi, and T. Tasaka. Sport Skill Classification Using Time Series Motion Picture Data. In *Proceedings of the 40th Annual Conference of the IEEE Industrial Electronics Society (IECON 2014)*, pages 5272–5277, Dallas (TX, USA), 2014.

[5] T. Maeda, A. Wakatani, and M. Yajima. Skill analysis using time series image frequency. In *Proceedings of International Workshop on Advanced Image Technology (IWAIT 2017)*, page (in USB stick memory), Penang (Malaysia), 2017.

[6] Y. Matsumoto. *Organization and Skill – Organization Theory of Preservation of Technique (in Japanese)*. Hakuto Shobo, 2003.

[7] Y. Mochizuki, R. Himeno, and K. Omura. Artificial skill and a new principle in sports (special issue on digital human : Measurement and modeling of human functions) (in japanese). *System, Control, and Information*, 46(8):498–505, 2002.

[8] S. Molau, M. Pitz, R. Schluter, and H. Ney. Computing mel-frequency cepstral coefficients on the power spectrum. In *Proceedings on IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'01)*, volume 1, pages 73–76. IEEE, 2001.

[9] T. Nomura, S. Oshikawa, Y. Suzuki, K. Kiyono, and P. Morasso. Modeling human postural sway using an intermittent control and hemodynamic perturbations. *Mathematical Biosciences*, 245, iss.1:86–95, 2013.

[10] T. Shiose, T. Sawaragi, K. Kawakami, and O. Katai. Technological scheme for the preservation of technique from ecological psychological approach (in japanese). *Ecological Psychology Research*, 1(1):11–18, 2004.

[11] P. Wardb and A. M. Williamsa. Do pattern recognition skills transfer across sports? a preliminary analysis. *Journal of Sports Sciences*, 22(2):205–213, 2004.