Acquisition of Embodied Knowledge on Sport Skill Using TAM Network

Isao Hayashi, Toshiyuki Maeda, Masanori Fujii,
Shuoyu Wang, and Tokio Tasaka

Abstract— **In this paper, we discuss sport technique evaluation using motion analysis model by neural networks and data mining methods. For students of university, we recorded the continuous forehand stroke of the table tennis in the video frames, and analyzed the trajectory pattern of nine marking points attached at subject's body with a coach's technique evaluation and the motion analysis model. As a result, we obtained some technique rules classified member of table tennis club, middle level player and beginner as fuzzy rules, and also estimated the movement of the marking points to improve in table tennis technique.**

I. INTRODUCTION

In human skill research for movement, the technique skill consists of hierarchical structure with a monofunctional layer to generate the single function result and a meta layer that adapted itself to an environmental change [1], [2]. In general, a skilled worker having high technique in a company acquires a hierarchical skill structure as internal model and they decide an action process by the internal model [2]. However, it is difficult to completely understand own the internal model. For an expert worker, typically, they observe own action representation objectively and achieve a high technique skill after fine-tuning internal model. According to this interpretation, we should make the internal model high-accuracy through two kinds of processes, which are the bottom-up processing of the intention expression to the representation action from the monofunctional layer and the top-down processing of the adjustment to the monofunctional layer from external observation, and we achieve a high technique skill as a result.

On the other hand, in the research of the skill movement of sport, many methods to use physical structured model and frame structured model by a movement analysis and a physiologic measurement analysis have been proposed [3]–[6]. In the paper [3], Mochizuki and et al. define a skill reproduced on an artifact as "artificial skill". They have proposed physical structured model with three-dimensional movement measurement technique by DLT(Direct Linear Transformation) method and estimate the mechanism of the most suitable throw movement of the pitcher of professional baseball. In addition, Kasai et al. applied the DLT method to forehand movement of table tennis, and observed the trajectory of the body movement by three dimensional analysis for producing a textbook suitable for beginners. Miyaki et al. [6] discuss experiential "use of motion-dependent forces" of the forehand stroke of table tennis using movement analysis.

In this paper, we discuss sport technique evaluation using motion analysis model by neural networks and data mining methods. Especially, we take the forehand stroke of the table tennis [5], [6] as an example of sports movement, and identify internal model with the neural network without using physical structured model and frame structured model. Perl et al. [7] employ Kohonen Feature Map as a neural network for analysis of the table tennis movement and estimate strategic structure of table tennis from analyzing the trajectory of the ball. In this paper, we recorded the continuous forehand stroke of the table tennis in the video, and identify an internal model with TAM network [8] and data mining methods from evaluation player's technique with three classes of expert player, middle player, and beginner. At first we selected total 15 subjects of university students who are seven expert players of the table tennis club, three middle players experienced table tennis, and five beginners, and recorded a trajectory pattern of forehand stroke for batting a ball with their the-handshake-grip racket by a high-speed camera. In addition, we constituted the observed data set from position coordinate and its speed of time-series data at nine marking points of the right upper arm by nine subjects with a coach's technique evaluation of three level. Next, we discussed similarity and difference between subjects from skill level of their forehand strokes by statistical analysis. Furthermore, we identified an internal model by TAM network, C4.5, Native Bayes Tree, and Random Forest, and discussed the relationship between the monofunctional layer and the meta layer in the internal model. Finally, we obtained some technique rules as fuzzy rules, and estimated the movement of the marking points to improve in table tennis technique.

II. ANALYSIS OF FOREHAND STROKE OF TABLE TENNIS

In the study of the movement analysis of sport, in general, the physical structure and the frame structure is clarified by electromyography which recorded action potential when muscular fiber was excited by the needle electrode, and

Isao Hayashi is with the Faculty of Informatics, Kansai University, Takatsuki, Osaka 569-1095, Japan (email: ihaya@cbii.kutc.kansai-u.ac.jp).
Toshiyuki Maeda is with the Faculty of Management Information, Hannan University, Matsubara, Osaka 580-0033, Japan (email: maechan@hannan-u.ac.jp).
Masanori Fujii is with the Faculty of Management Information, Hannan University, Matsubara, Osaka 580-0033, Japan (email: fujii@hannan-u.ac.jp).
Shuoyu Wang is with the Graduate School of Engineering, Kochi University of Technology, Kami, Kochi 782-8502, Japan (email: wang.shuoyu@kochi-tech.ac.jp).
Tokio Tasaka is with the Faculty of Health and Sports Science, Doshisha University, Kyotanabe, Kyoto 610-0394, Japan (email: ttasaka@mail.doshisha.ac.jp).

marking observation methods by marking points attached body to detect X-Y position and speed. However, in this paper, we suppose that the technique skill consists of hierarchical structure with a monofunctional layer to generate the single function result and a meta layer that adapted itself to an environmental change, and analyze the table tennis skill from trajectory data of forehand stroke and coach's technique evaluation for the skill using TAM network. Figure 1 shows the structure of the proposed internal model.

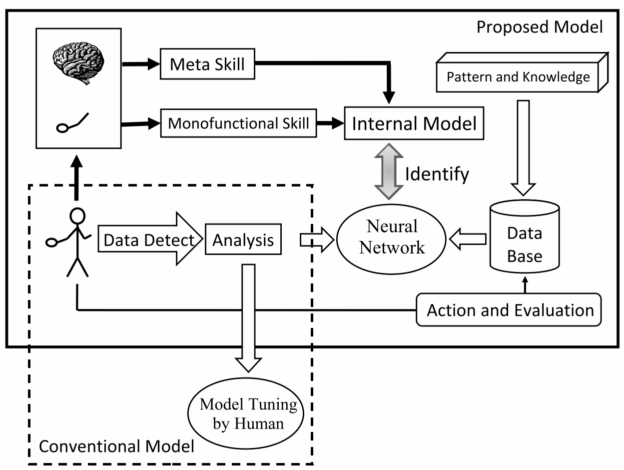

Fig. 1. Proposed System

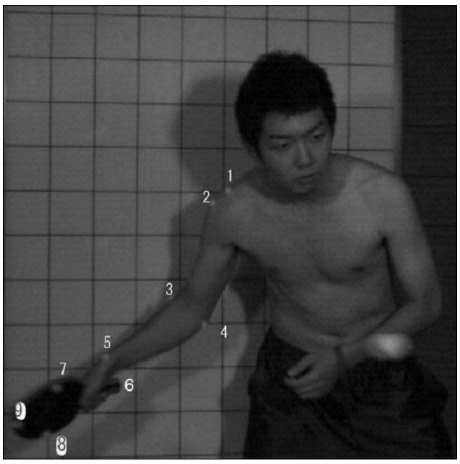

Fig. 2. Mesurement Markings

In the experimental measurement, we selected 15 students of Hannan University as subjects. Fifteen people of the subject are seven students belong to the Hannan University table tennis club as expert player of observation data, three students who have been table tennis club members in a junior high school and high school as middle player, and five students as the beginner without the experience of the table tennis.

For the experiment attempts, we set nine marking points for observation data on the right upper arm of the subject, which are 1)the acromioclavicular joint, 2)the acromion, 3)the head of radius, 4)the head of ulna, 5)the styloid process of radius, 6)the styloid process of ulna, 7)the right apex point in the racket edge, 8)the left apex point in the racket edge, and 9)the upper apex point in the racket. Figure 2 shows the setting position of measurement marking.

A pitches machine (Yamato table tennis Co., Ltd., TSP52050) were set at about 30cm distance from the end line of the table diagonally in the extended line of subject, and a ball was distributed to throw at elevation of 20 degrees, 25 speed levels, and 30 pace levels. The subject have to return a ball which bounded in the 75cm inside from the end of the table to the opposite side in the forehand cross. For the observation of the movement trajectory of the ball, we used a high-speed camera (Digimo Company, VCC-H300, resolution: $512 \times 512 pixel$, frame rate: $90 fps$) placed in front $360cm$ of the subject and $130cm$ in height. While a subject return a ball, we recorded his forehand stroke in the video for 10 minutes(Figure 3).

Fig. 3. Pictures of Subject

We extracted still images of 40 frames to 120 frames from video memory that a subject swung a racket from taking back to the end of swinging. In each frame image, we calculated the two-dimensional (x, y) coordinate of nine marking points as the original position at the shoulder of the subject of the first frame. As an example, we show the observation position of marking points in two-dimensional coordinate, and the speed of the horizontal direction (x) of two expert players, two middle players, and three beginners in Figure 4 and Figure 5. In addition, we show the minimum and the maximum value of the coordinate position of horizontal direction (x) at the first marking (M_1), the fourth marking (M_4), and the ninth marking (M_9) of Figure 4 in Table I.

From Figure 4, Figure 5 and Table I, the following result are obtained.

- By comparison with two expert players, the coordinate of the position from M_1 to M_9 was fitted extremely. The

TABLE I
MIN AND MAX POSITION OF X-DIRECTION OF MARKINGS

	M_1		M_4		M_9	
	Min	Max	Min	Max	Min	Max
Expert	-3	114	-29	254	-267	372
Middle	-10	116	-25	236	-218	577
Beginner	-33	152	-50	239	-214	697

correlation coefficient was obtained as $x = 0.985, y = 0.790$. Therefore, the expert player acquires a common expertise ability skill to swing a racket by similar form. In addition, the expert player return a ball by a smooth common forehand drive because the trajectory draws an oval shape without the fluctuation.

- From the data of the expert player, the speed of the moment hitting a ball was maximum at all observation marking points. They swing forehand throw (positive speed) from take back (negative speed) smoothly. In other words, they acquire a technique skill to be the maximum speed in the impact hitting a ball.
- By comparison with two middle players, the coordinate of the position from M_1 to M_9 was partly fitted. The correlation coefficient was $x = 0.919, y = 0.607$. The middle player acquires an expertise skill well because the trajectory of swing resemble each other. However, the trajectory doesn't draws an oval smooth forehand drive.
- From the data of the middle player, the speed of M_7 and M_9 becomes the two peaks form. We should notice that they have adjusted speed at the moment of the impact to hit a ball with the racket.
- By comparison with three beginners, the coordinate of the position from M_1 to M_9 was quite different. The correlation coefficient was $x = 0.073, y = -0.04$. The position coordinates in M_1 are largely between each beginners, and the width of the trajectory is in particular big. The beginner shoulder shakes in comparison with the expert player and the middle player. In addition, the position coordinate of M_7 and M_9 is quite different each other. From these results, how to swing a racket by the beginner is variety.
- From the speed data of the beginner, they stopped speed just before hitting a ball and waited a ball to coming in M_3 to M_9, and so-called, "a movement to go to meet a ball by a racket" was seen. In comparison with the expert player and middle player in M_1, the shoulder of beginners shakes greatly. Furthermore, "a movement to delay body" is seen, that is a shoulder and an elbow move too much for the movement of the racket because the speed in M_1 and M_4 is observed even at the frames the speed of M_7 and M_9 is zero.
- From Table I, the expert player swings a racket compactly small in the horizontal direction. The expert player acquires a technique skill to maximize speed just at a moment hitting a ball from Figure 5. Though the beginner swings the big width in the horizontal direction, they adjust the speed of the racket swing before an impact and the movement to delay body was detected. The middle player has a middle technique skill between the expert and the beginner.
- From all results, there is no category with the same technique pattern as for the beginner. On the other hand, the expert player and middle player constitute the category with the same technique level.

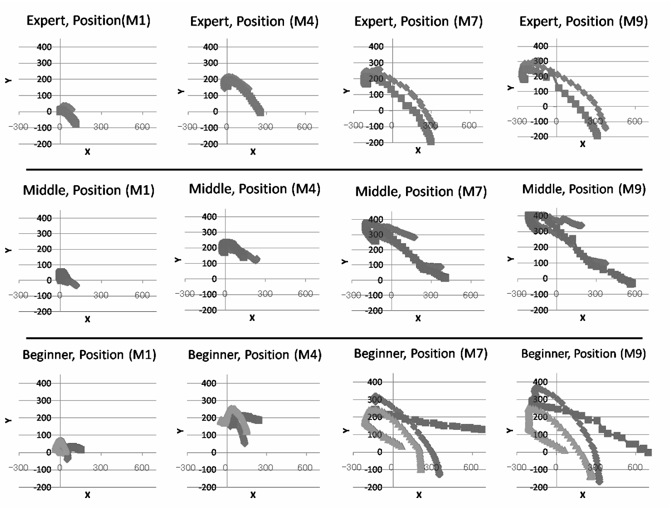

Fig. 4. Position of Markings

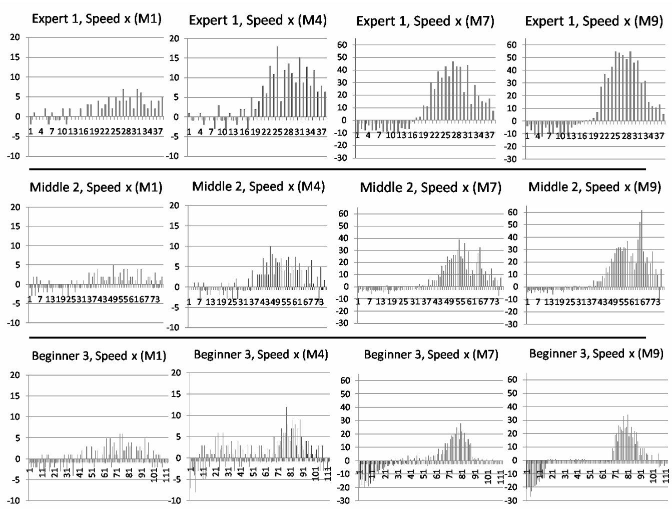

Fig. 5. Speed of Markings

III. IDENTIFY THE INTERNAL MODEL USING TAM NETWORK

A Topographic Attentive Mapping (TAM) network is based on a biologically-inspired model constructed in imitation of the human vision system. The network structure consists of four layers: the feature layer, the basis layer, the category layer, and the class layer. If the network produces inaccurate output, the attentional top-down signal modulates the synaptic weight in the class and basis layers in order to minimize the difference between the output and the supervised data by a winner-takes-all algorithm. Simultaneously, a node is added to the category layer until the output accuracy is improved. The structure of the TAM network is shown in Figure 6.

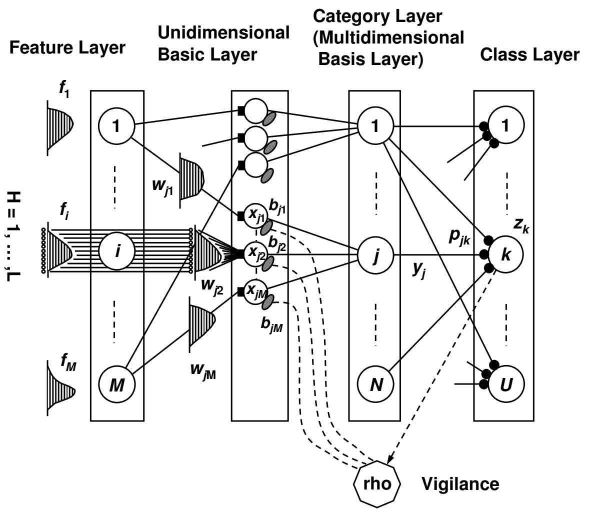

Fig. 6. TAM Network

The activity value x_{ji} of each node of the unidimensional basis layer is calculated by the distributed synapse weight w_{jih} between the feature layer and the inhibitory synapse weight b_{ji} by the vigilance parameter ρ between the class layer. Output y_j from the category node to the class layer is calculated as follows:

$$y_j = \prod_{i=1}^{M} x_{ji} = \prod_{i=1}^{M} \frac{\sum_{h=1}^{L} f_{ih} w_{jih}}{1 + \rho^2 b_{ji}}. \qquad (1)$$

In the class layer, the maximum value of each node output is adopted as the output of the TAM network.

$$K = \{k | \max_k z_k\} = \{k | \max_k \sum_{j=1}^{N} y_j p_{jk}\} \qquad (2)$$

where $p_{jk},\ k = 1, 2, \cdots, U$ is the synapse weight between a class node and a category node.

Now, let K^* denote the "correct" supervised output. If the output K of the TAM network does not correspond with the supervised output class K^*, the "attention" mechanism is invoked, and the vigilance parameter ρ increases to the subject level of $z_{K^*}/z_K \geq OC$ or the maximal vigilance level $\rho^{(max)}$, where OC is the threshold.

If $z_{K^*}/z_K < OC$ then repeat
(a) $\rho = \rho + \rho^{(step)}$
(b) equation (1) and (2)
until either $z_{K^*}/z_K \geq OC$ or $\rho \geq \rho^{(max)}$.

When the vigilance parameter ρ reaches its maximum level $\rho^{(max)}$, one new node is added to the category layer. When the condition of $z_{K^*}/z_K \geq OC$ is satisfied in the interactive processing of Equation (3), the learning process of synapses is started and learning parameters of $w_{jih},\ p_{jk}$ and b_{ji} are updated.

We identified the internal model of subject using TAM network. The technique skill of the table tennis depends on the time-series data of position coordinate of the observation marking. Therefore we constituted the observed data overlapping with input data of five consecutive frames data from the second frames to the sixth frames later in each frame data for nine subjects. The skill evaluation as output value is three classes of the expert player, the middle player, and the beginner. The observation data consist of 90 input variables, and the three classes output since a position of each observation marking is a two-dimensional coordinate of (x, y).

The training data (TRD) is constituted with two expert players, two middle players, and three beginners, and the checking data (CHD) is constituted with one expert player and one beginner. The result strongly depends on which subject data is used for learning or evaluation. Therefore we calculated the correlation coefficient of the position coordinate at each marking, and constituted three kinds of the observed data which are the data set A divided high two subjects of the correlation coefficient for TRD and CHD among four beginners, and the data set B and C divided high two subjects of the correlation coefficient for TRD among four beginners. In addition, for the expert player, we assign high two subjects of the correlation coefficient for TRD and remaining one subject for CHD.

TABLE II
RECOGNITION RATE OF DATA SETS

	Recognition Rate(%)		
	TRD	CHD	Ave.
Data Set A	53.7	57.5	55.6
Data Set B	56.9	43.3	50.2
Data Set C	55.2	42.3	48.8

The result of TAM network is shown in Table II. In each data set, the recognition rate for TRD and CHD is not so high. It is thought that this reason has a big different to the number of the observation data of each class. Therefore, for data set A which average recognition had best, we constituted the observed data overlapping with input data of five consecutive frames data from next frames to fifth frames later in each frame data, and we let the number of observed data increase. The result is shown in Table III. TAM(A) is recognition rate for data set A, and TAM($A+$) shows recognition rate for the revised data set A which performed data correction between classes. The result with C4.5 which is data mining method, Native Bayes Tree(NBT), Random

Forest(RF) show at the same time, where the result of the data mining methods is recognition rate for data set A.

TABLE III

RECOGNITION RATE OF REVISED DATA SETS

	Recoginiton Rate(%)		
	TRD	CHD	Ave.
TAM(A+)	61.2	43.0	52.1
TAM(A)	53.7	57.5	55.6
C4.5	98.1	43.3	70.7
NBT	100.0	32.8	66.4
RF	100.0	25.4	62.7

We should notice that the recognition rate of the TAM network for data set $A+$ improves than A from these results. On the other hand, the recognition rate for TRD of NBT and RF is provided with 100%, and we thought it with overlearning for TRD. The recognition rate for CHD is extremely bad. The C4.5 showed a high recognition result for TRD and CHD. The recognition rate of the TAM network after revised data correction between classes showed a result at the same level for CHD with C4.5.

TABLE IV

SENSITIVITY OF INPUT VARIABLES

Number of Input Var.	Omitted Input Var. and Recognition Rate(%)				Selected Input Variables
	M_1, M_2	M_3, M_4	M_5, M_6	M_7-M_9	
18	-	-	-	-	-
12-14	42.9	57.4	51.1	48.2	M_1, M_2
8-10	-	45.9	48.4	41.6	M_7-M_9
4	-	42.9	42.0	-	M_5, M_6
-	-	-	-	-	M_3, M_4

Next, we analyzed the sensitivity of the variable of the marking points with TAM network. We obtained the priority of marking points with the data set($A+$) for 18 input (90 input variables in the data set) of nine points of markings and one output of three classes by the TAM network. We removed four input variables (20 input variables) from 18 input variables (90 input variables) temporarily, and obtained the input variables that recognition rate was the lowest. The input variable that recognition rate was the lowest represents the highest priority since the recognition rate decreases by removing the input variable.

The result of the sensitivity analysis is shown in Table IV. When M_1 and M_2 are temporarily removed, the recognition rate of the TAM network was the 42.9% and that was the lowest recognition rate. Therefore, the input variables as the first priority are M_1 and M_2. As a result, the important input variables were obtained in order of $M_1, M_2 \to M_7, M_8, M_9 \to M_5, M_6 \to M_3, M_4$. We should notice that the recognition rate is continuously getting to down in M_1, M_2 and M_7, M_8, M_9. On the other hand, the recognition rate is increasing when M_5, M_6 and M_3, M_4 were removed. From these results, the important marking points to distinguish the expert player, the middle player, and the beginner are obviously 1)the acromioclavicular joint, 2)the acromion, and 7) to 9) in the racket. This result is consistent with an analysis conclusion in Figure 4 and Figure 5.

Now, we express recognition rate calculated with the i-th input variable as R_i. We define the importance of the input variable using R_i as the following P_i.

$$P_i = \frac{R_i - R_{i-1}}{\sum_i |R_i - R_{i-1}|} \tag{3}$$

Since the value of P_i express the ratio of the deviation of the recognition rate between the i-th variable and the $i-1$-th variable in the total deviations, the positive of P_i means the evaluation value that can distinguish between classes, and the negative means the evaluation value to express the similarity between classees. In Table IV, $P_{M_1,M_2} = 0.88$, $P_{M_7-M_9} = 0.06$, $P_{M_5,M_6} = -0.02$, $P_{M_3,M_4} = -0.04$ are obtained. Figure 7 shows the value of P_i of the marking point.

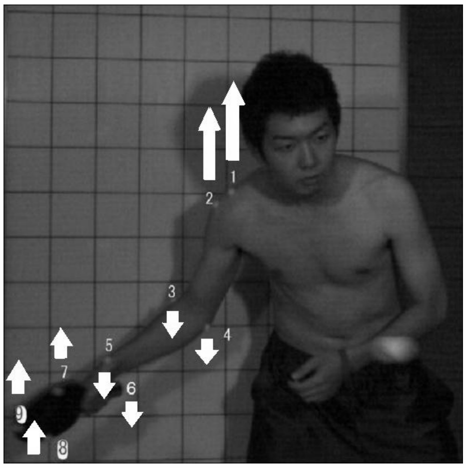

Fig. 7. Priority of Markings

Finally we acquired the technique skill as the fuzzy rule. The TAM network consists of four layers of hierarchical structure. The feature layer and the basis layer of the lower level represent the monofunctional concept, and the category layer, and the class layer of upper level represent the meta concept. Therefore, we can acquire the relationship between the monofunctional skill and the meta skill with the fuzzy rule.

We selected first the J-th category node where p_{jk} became the maximum in each class node of the expert player, the middle player, and the beginner for data set ($A+$), and next calculated w_{Ji} of a category node of the J-th category node every input variable. As a result, we acquired the

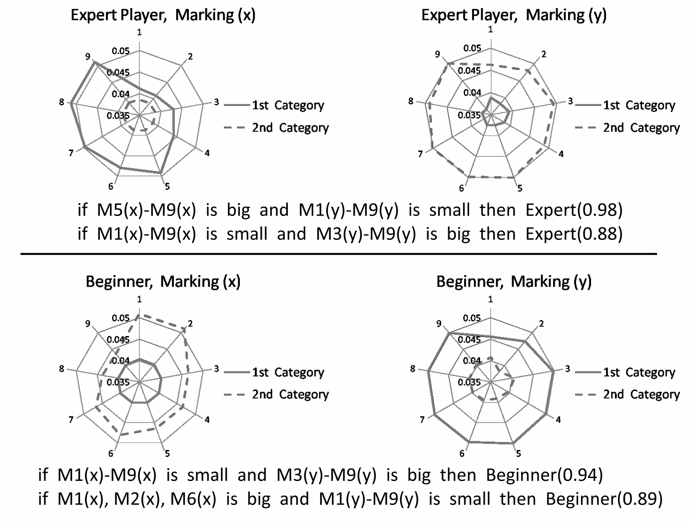

Fig. 8. Rule of Technique Skill

monofunctional skill and the meta skill as fuzzy rule format.

$$w_{Ji} = \frac{\sum_{h=1}^{L} w_{Jih}}{L}, \quad for\ \forall i \tag{4}$$

$$J = \{j | \max_{j} p_{jk}, \quad k = 1, 2, 3\} \tag{5}$$

The result is shown in Figure 8. A rule of a meta skill is acquired for the expert player and the beginner.

IV. Conclusion

In this paper, we evaluated technique skill of a forehand stroke of table tennis with three classes, and identified the internal model of a technique skill with TAM network. In addition, we discuss the monofunctional skill and the meta skill to improve technique of table tennis.

References

[1] T.Shiose, T.Sawaragi, K.Kawakami, and O.Katai: Technological Scheme of Skill Succession from Ecological Psychological Approach, *Journal of Japan Society for Ecological Psychlogy*, Vol.1, No.1, pp.11-18 (2004)

[2] Y.Matsumoto: *Skill and Organization*, Hakuto (2003)

[3] Y.Mochizuki, R.Himeno, and K.Omura: Artificial Skill and a New Principle in Sports, *System, Control and Information*, Vol.46, No.8, pp.498-505 (2002)

[4] M.Suwa: A Cognitive Model of Acquiring Embodied Expertise Through Meta-cognitive Verbalization, *Transactions of the Japanese Society for Artificial Intelligence*, Vol.23, No.3, pp.141-150 (2008)

[5] J.Kasai, T.Mori, M.Yoshimura, A.Ota: A Study of the Forehand Stroke of Table Tennis by Three Dimensional Analyses Using DTL Method, *Waseda Journal of Human Sciences*, Vol.7, No.1, pp.119-127 (1994)

[6] M.Miyaki, N.Ashida, N.Takashima, T.Azuma, H.Tsuruta: Movement Analysis of Forehand Stroke in Table Tennis, *The 42th Conference of Japan Society of Physical Education, Health and Sport Sciences*, pp.681 (1991)

[7] J.Perl, A.Baca: Application of Neural Networks to Analyze Performance in Sports, *Proceedings of the 8th Annual Congress of the European College of Sport Science*, in Salzburg (2003).

[8] I.Hayashi and J.R.Williamson: Acquisition of Fuzzy Knowledge from Topographic Mixture Networks with Attentional Feedback, *the International Joint Conference on Neural Networks (IJCNN '01)*, pp.1386-1391, in Washington DC (2001)