

Imitation and Primitive Symbol Acquisition of Humanoids by the Integrated Mimesis Loop

Tetsunari Inamura^{*1*2} Yoshihiko Nakamura^{*1*2} Hideaki Ezaki^{*3} Iwaki Toshima^{*1}

^{*1} Dept. of Mechano-Informatics, The University of Tokyo, JAPAN

^{*2} Japan Science and Technology Corporation, CREST program, JAPAN

^{*3} Kawasaki Heavy Industries Co., Ltd., JAPAN

Abstract

Mimesis is a primitive learning framework and origins of human intelligence. We have developed a behavior acquisition and understanding system based on the mimesis. This system is able to abstract observed others' behaviors into conceptual symbols, to recognize others' behavior using the primitive symbols, and to generate self motion patterns using the primitive symbols. In this paper, we mention the integration of mimesis loop which is the acquisition and development system based on mimesis, and confirmation of the feasibility against whole body motions on virtual humanoids.

1 Introduction

Designing of the behavior of humanoid robots have many difficulties. Many methods[1] are proposed in order to overcome the problems. For example, a method which compensates ZMP conditions[2], and a learning method which uses heuristic evaluation functions, were the main currents. Recently, contrary to such approaches in which developers give the robots control systems, behavior acquisition methods using imitation of human's behavior have attracted a great deal of attention[3][4]. Indeed, a module referred as "Mirror neuron" has discovered in the field of neuro physiology[5] which fires when a human observes a specific behavior or when a human intend to act the specific behavior. For the facts, a hypothesis is proposed that the mirror neuron is relevant to the imitation learning, and the imitation function is provided with animals.

Here, we touch on a final image of our research on humanoid robots. They are communication function between human with natural language, planing function for complex behavior using symbols given by

human, and so on. For such purpose, it is necessary to use meta-level symbolized instructions like "Walk toward there" or "Take that target".

When the mirror neuron has relevancy to the symbols, it is possible for humanoid robots not only to generate its own behavior from observation of others' behavior but also to emerge symbols from the observation and to generate behavior from the instructed symbols. It namely has a possibility unification of the language functions and the behavior control functions.

This concept have relevance to *Mimesis*[6] which is one of the framework models in the filed of cognitive psychology. In the mimesis frameworks, human creates certain symbols from observation and understanding of others' behavior, then the human generates self behavior and acquires natural languages. Thus it is regarded as the origin of human's intelligence.

In this paper, we focus on *Mimesis* on the presupposition that the mimesis is valid for imitation learning. We mention the system integration based on the mimesis in which abstraction of observed others' behavior into symbols and generation of natural behavior from the abstracted symbols, are realized. The symbols in the system have a possibility that it enables the robots to communicate with human using language, and to acquire language through everyday behavior.

In section 2, we define the basic module called as self motion elements, and the abstracted symbols called as primitive symbols. In section 3, we mention the method of abstraction from observed others' motions, and in section 4, we mention the method of reproduction of whole body motion. In section 5, we investigate the validity of the method through motion generation experiments on a virtual humanoid robot. We conclude in section 6.

2 Mimesis loop using self motion elements and primitive symbols

2.1 Generation, abstraction, and recognition of whole body behavior

In feedback error learning proposed by Kawato et al[7], dynamics in a observed behavior is acquired and represented as a feed forward model. This model is able to abstract the dynamics, however, the consideration for symbolization is not well studied, therefore it is not able to be relevant to high order information processing such as recognition of behaviors. Schaal[3] et al have proposed a imitation learning method in which some abstracted modules referred as *primitive behavior* are prepared by developer in order to combine a complex behavior. There are some suggestions that this kind of modules is needed for the symbolization.

In Kuniyoshi's learning approach[8], the rule of sequence of behaviors is constructed by observation and symbolization of human's demonstrations. The approach can reproduce complex behaviors with the abstraction and symbolization, however, it is difficult to be applied to humanoid robots from lack of dynamics point of view.

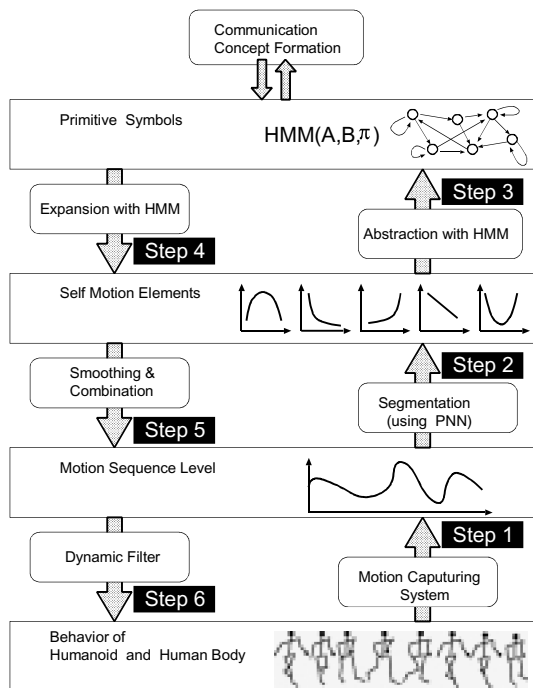


Fig. 1 : The mimesis loop for imitation learning

Here, we propose a model which can abstract whole body motion into a certain symbols, generate motion

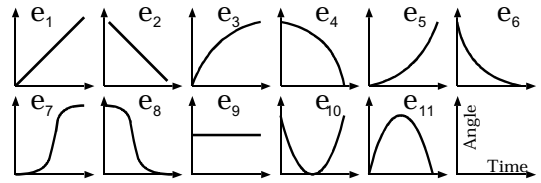


Fig. 2 : Self motion elements

patterns from the symbols, and recognize total behavior. The model is referred to as mimesis-loop with the precondition of the mimesis, as shown in Fig.1. We can reappear the mirror neuron explained before from a engineering point of view, when the dynamics is abstracted and represented as the symbols. In this paper, we introduce “primitive symbols” as this type of symbols, and “self motion elements” as basic motion patterns which compose the primitive symbols.

Mimesis loop consists of two phases. In the first half, observed motions are transformed into self motion elements by comparing, and are abstracted as primitive symbols. Observed motion patterns are analyzed into the self motion elements, and the sequence of the motion elements are abstracted into primitive symbols, regarded as a series of behavior. We adopted Hidden Markov Models (HMMs) for the description of the relation between sequence of motion patterns and primitive symbols.

HMM is one of stochastic processes which takes time series data as an input, then outputs probability that the data is generated by the HMM. This probability is referred to as likelihood, which enables the mimesis-loop to recognize others' behavior, and to generate humanoids' behavior, and to create primitive symbols. The detail application of the HMMs is explained in the next chapter.

2.2 Definition of Self Motion Elements

Self motion elements mean the simple and primitive motion patterns like “moving the hand slowly” and “thrashing one's legs”. To put it concretely, it is sequential joint angle pattern of each joint.

In this paper, we adopted 11 basic motion patterns

$$\mathbf{e} = \{e_1, e_2, \dots, e_{11}\} \quad (1)$$

as the self motion elements, as shown in Fig.2. The horizontal axis indicates time, and vertical axis indicates joint angles. The self motion elements are partial piece of one joint motion. A whole body motion consists of the combination of the self motion elements.

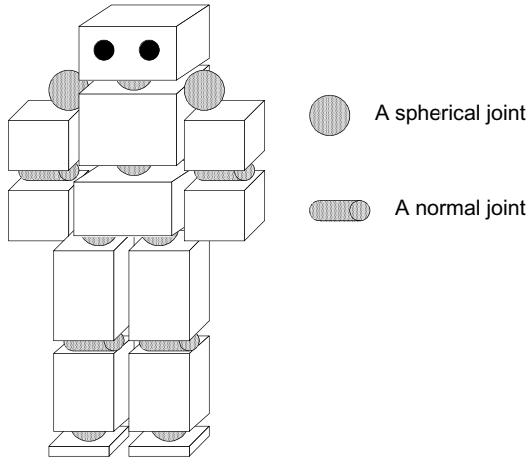


Fig. 3 : Configuration of degree of freedom

The virtual humanoid used in this research has 28 DOFs as shown in Fig.3. A normal joint corresponds to one self motion elements, and a spherical joint corresponds to three self motion elements.

2.3 Definition of Primitive Symbols

When a dynamics included in the element combination is abstracted and represented as certain symbols, it is possible for the robot to recognize the others' motion and to generate whole body motions from symbols. We refer to the symbol as *primitive symbols*.

3 Abstraction of the whole body motion using primitive symbols

The abstraction of primitive symbols consists of two phase. At first phase, the robot observes others' motion and analyzes the motion into self motion elements. At second phase, the dynamics in the sequence elements are abstracted as primitive symbols using HMMs.

3.1 Analyzing observed motion into self motion elements

Properly speaking, it is better to analyze others' motion in the camera coordinate system into self motion elements in the body coordinate system. In this paper, we omitted the coordinate transformation process with the assumption that the robot observes others' motion in the absolute coordinate system.

At first, motion patterns are analyzed into motion segments using a border time on which the joint angle has big acceleration. Next, this segments are recognized as the self motion elements using Probabilistic Neural Network (PNN)[9]. PNN is a neural network model which uses radial basis function. This model classifies input patterns into prepared classes based on each feature vector. It has two advantage that it is tolerant toward noise, and it outputs the result with probabilities of the sureness of classification.

3.2 Abstraction of motion sequence using HMMs

Next, a HMM which can output the sequence of self motion elements plausibly, is generated. (corresponds to Step 3 in Fig.1) HMMs consists with three parameters; (1)transition probability matrix A_{ij} which means the probability of transition from status i to status j , (2)output probability matrix B_{ix} which means the probability of output x at the status i , (3) initial distribution probability π . The sequence of self motion elements correspond to the sequence of output symbol x . The primitive symbols are defined by the set of these three parameters as following:

$$P_s \stackrel{\text{def}}{=} \{A, B, \pi\} \quad (2)$$

The three parameters are calculated by Baum-Welch method in order for the system to output plausible output sequence.

Let n the number of self motion elements pieces, which is the analysis result of the observed motion sequence. In this case, the generated HMM has n states, the size of transition probability matrix A_{ij} is $n \times n$, and the size of B_{ix} and π is $1 \times n$.

In the case of an unknown motion input, the robot creates a new HMM. However, when an already-known motion is inputted, the robot must recognize the motion and output the result as the primitive symbols. HMMs are originally used for the pattern recognition, therefore, this kind of applications are easy to be realized.

4 Generation of whole body motion from primitive symbols

The generation of whole body motion consists with following two procedures.

1. Generating sequence of self motion elements from primitive symbols (Step 4 in Fig.1)

2. Transforming the self motion elements into joint angles, and considers dynamics conditions[10] (Step 5,6 in Fig.1)

4.1 Generation of sequence of self motion elements using HMMs

The target of this phase is to generate symbol sequence from HMMs and to transform them into self motion elements. When let the symbol sequence \mathbf{O} as following,

$$\mathbf{O} = o_1, o_2, \dots, o_T \quad (3)$$

$$o_i \in \mathbf{e} \quad (4)$$

the likelihood $P(\mathbf{O}|P_s)$ is used for description of the relation between motion patterns and the primitive symbol. Therefore, the generation is relevant to search a motion pattern which has the best likelihood among the entire patterns of \mathbf{O} .

There is an example of generating patterns using continuous mixture HMMs in speech synthesis field, however, it depends on the left-to-right models which is often used in speech recognition systems. The left-to-right models is not suitable for periodic motion such as human's behavior. Thus we propose a pattern generation method using normal discrete HMMs

Most simple way to search the best pattern is to scan the whole pattern space of the \mathbf{O} , and find the maximum likelihood. However, it is difficult to adopt this method because the size of the search space is 11^T when the length of motion sequence is T . Thus we have adopted generic algorithm (GA) for the searching.

In order to encode a motion pattern into a chromosome, T genes which correspond to 11 self motion elements are used, on the assumption that the number of motion sequence outputted by the HMMs is T . As the fitness of the chromosome, the likelihood of that the motion patterns are generated by the HMM, $P(\mathbf{O}|P_s)$ is used. It have also adopted translocatoin not simple crossover and mutation. It is suitable for the evolution to keep a series of behavior because the block of self motion elements indicates the series of behavior.

4.2 Motion modification using dynamics filter

The time series data of all joint angles, namely the kinematic whole body motion of the humanoid is decided by the combination of outputted self motion elements from the HMMs. However, there is no guarantee that the whole body motion satisfies dynamics conditions because of the discontinuity of the whole body

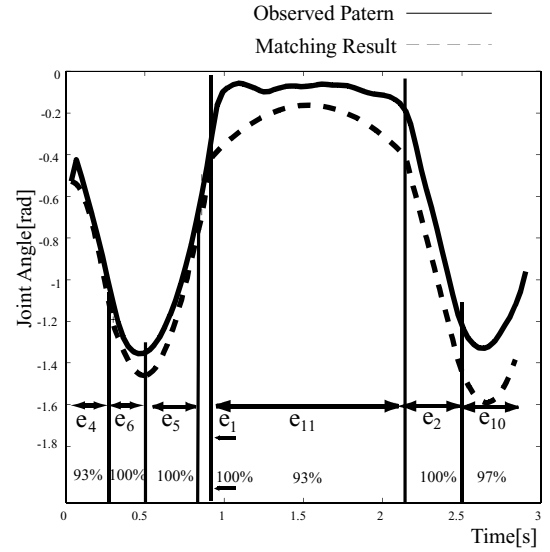


Fig. 4 : Segmented motion elements(left knee)

motion at the joint of motion elements. Thus, transforms the outputted motion into appropriate motion to be satisfy the dynamics conditions using *Dynamics Filter* [10]. (This process is corresponds to Step 6 in Fig.1.)

5 Simulation of generating whole body motion

5.1 Simulation of motion abstraction

For the verification of the method, human's motion data have observed via motion capturing system. The capturing system can measures 28 joint angles of all DOFs with sampling time 33[ms]. Four walking data; person A, person B, tired man, and old man, have observed for verification experiments.

The captured walking motions are shown in from Fig.5 to Fig.8 From here, focusing on the angle of left knee, the time series of angle is as solid line shown in Fig.4. The horizontal axis indicates time [sec], and the vertical axis indicates joint angle [rad]. The graph shows one and half cycle of the periodic walking motion. The circled number for each segment is the self motion elements, and the percentage is probability of PNN matching. The broken line indicates the angle time series constructed by the sequence of self motion elements, $\{e_4, e_6, e_5, e_1, e_{11}, e_2, e_{10}\}$.

Primitive symbols used in recognition	Input Behavior			
	Testee A	Testee B	Tired testee	Old testee
$P_s^{(a)}$	1.2×10^{-1}	5.4×10^{-31}	3.4×10^{-46}	6.1×10^{-18}
$P_s^{(b)}$	1.2×10^{-19}	2.8×10^{-8}	3.6×10^{-43}	4.8×10^{-14}
$P_s^{(t)}$	2.3×10^{-21}	5.4×10^{-32}	2.4×10^{-10}	8.7×10^{-14}
$P_s^{(o)}$	4.6×10^{-25}	5.4×10^{-42}	2.4×10^{-48}	2.3×10^{-1}

Table 1 : Recognition result of others' motion using HMMs.

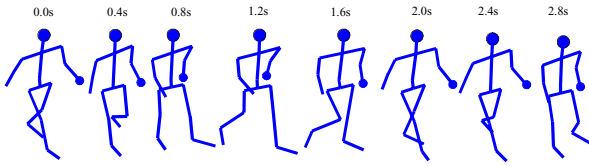


Fig. 5 : A walking motion of testee A

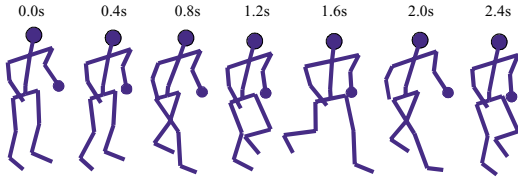


Fig. 6 : A walking motion of testee B

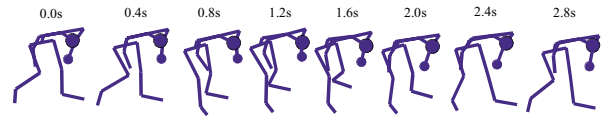


Fig. 7 : A walking motion when the testee was tired



Fig. 8 : A walking motion by an old man

5.2 Simulation of motion generation

We have practiced an experiment in which most plausible sequence of self motion elements is generated from the three parameter A, B, π which corresponds to primitive symbols. The condition of the GA is that the number of gene is 7, the number of chromosomes is 500, the crossover ratio is 0.9, the mutation ratio is 0.95, and elite selection was adopted. At about 500th generation, a sequence of self motion elements as same as shown in Fig.4 is generated. A calculation time is about 3[sec] on Pentium-III 600MHz, which is enough rapid to apply for motion generation based on symbolic communication between users.

A whole body motion transformed from the outputted sequence of self motion elements is shown in Fig.9. (This process corresponds to Step 5.) It is obvious that the generated motion pattern as same as the observed motion pattern, it also shown in Fig.4.

Fig.10 shows a motion which is output of the dynamics filter where the input motion is Fig.9. As the result indicates, that awkward motion has been translated into natural human like motion with long steps and modification of waist posture.

5.3 Recognition of others' motion

Here, we explain a method to recognize the observed motions using HMMs. Two primitive symbols, $P_s^{(a)} = \{A^{(a)}, B^{(a)}, \pi^{(a)}\}$ and $P_s^{(b)} = \{A^{(b)}, B^{(b)}, \pi^{(b)}\}$ was generated from observation of two walking behaviors by testee A and B. Another primitive symbols was generated from walking behavior when the testee was tired ($P_s^{(t)}$), and when the testee was an old man ($P_s^{(o)}$). After that, observes another walking behaviors of A, B, tired testee, and old testee, then executes the recognition process for each HMMs. The result is shown in Table 1. Each value indicates the likelihood of the motion generation. On the one hand, the likelihood of the walking pattern A generated by $P_s^{(a)}$ is 10^{17} times as large as the likelihood generated by $P_s^{(b)}$. On the other hand, the likelihood of the pattern B generated by $P_s^{(b)}$ is 10^{23} times as large as the likelihood generated by $P_s^{(a)}$. This result shows that the ratio of the likelihood is useful for the recognition. When an unknown behavior is inputted, it is possible to judge the novelty using certain threshold of the likelihood.

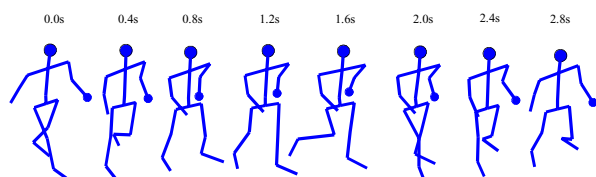


Fig. 9 : A generated motion

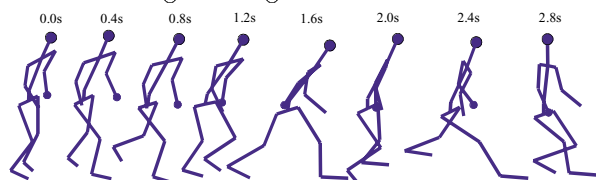


Fig. 10 : A motion after passing dynamics filter

6 Summary and Conclusions

We focused on mimesis learning and proposed mimesis-loop using self motion elements and primitive symbols. The feature of this system is that it is possible to realize following three function in the same model: (1)To abstract the dynamics of human's motion, (2)To generate natural motion patterns from the symbolized dynamics, and (3)To recognize others' behavior using primitive symbols. An abstraction method using PNN for analysis of the observed motion into self motion elements and HMMs for represent the motions' dynamics have explained. Also a generation method using HMMs and GA have explained.

Current difficulty is that a HMM represents not dynamics of whole body but dynamics for only one joint angle. There must be dynamics or correlative relationship between each joint, therefore it is not enough satisfied to express the whole body motion with leaving the current difficulty as it is. For the future works, we plan to introduce new self motion elements which can treat whole body motion. We also have a prospect for using the result after passing the dynamics filter as fitness of the GA, because it is suitable to generate natural motion from the HMMs.

In this paper, the functions of the proposed system are the abstraction of others' motion and the reproduction of self motion. We plan to expand the system in order to refine the self motion by inputting the output self motion into the observation process as an other's motion. Finally, the research aims to realize humanoids which can communicates with human and makes plans using symbols.

Acknowledgement

This research was partially supported by the "Robot Brain Project" of Core Research for Evolutional Science and Technology (CREST) of the Japan Science and Technology Corporation (JST).

References

- [1] Ken'ichirou Nagasaka, Masayuki Inaba, and Hirochika Inoue. Dynamic Walking Pattern Generation for a Humanoid Robot Based on Optimal Gradient Method. In *Proc. of the 1999 IEEE International Conference on Systems, Man, and Cybernetics*, pp. FA22-3, 1999.
- [2] Jong H. Park and Yong K. Rhee. ZMP Trajectory Generation for Reduced Trunk Motions of Biped Robots. In *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robotics and Systems (IROS'98)*, pp. 90-95, 1998.
- [3] Stefan Schaal. Is imitation learning the way to humanoid robots? *Trends in Cognitive Sciences*, Vol. 3, No. 6, pp. 233-242, 1999.
- [4] Maja J Mataric. Getting humanoids to move and imitate. *IEEE Intelligent Systems*, pp. 18-24, 2000.
- [5] V. Gallese and A. Goldman. Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, Vol. 2, No. 12, pp. 493-501, 1998.
- [6] Merlin Donald. *Origins of the Modern Mind*. Harvard University Press, Cambridge, 1991.
- [7] M. Kawato, K. Furukawa, and R. Suzuki. A hierarchical neural-network model for control and learning of voluntary movement. *Biological Cybernetics*, Vol. 57, No. ?, pp. 169-185, 1987.
- [8] Yasuo Kuniyoshi, Masayuki Inaba, and Hirochika Inoue. Learning by Watching: Extracting Reusable Task Knowledge from Visual Observation of Human Performance. *IEEE Transaction on Robotics and Automation*, Vol. 10, No. 6, pp. 799-822, 1994.
- [9] Philip D. Wasserman. *Advanced Methods in Neural Computing*. Van Nostrand Reinhold, 1993.
- [10] Katsu Yamada and Yoshihiko Namamura. Dynamics filter — concept and implementation — on-line motion generator for human figure —. In *International Conference on Robotics and Automation*, pp. 688-695, 2000.