Proceedings of the 2002 IEEE
International Conference on Robotics & Automation
Washington, DC • May 2002

Acquisition and Embodiment of Motion Elements in Closed Mimesis Loop

Tetsunari Inamura*1*2 Iwaki Toshima*2 Yoshihiko Nakamura*2*1
*1 The CREST program, Japan Science and Technology Corporation, JAPAN
*2 Dept. of Mechano-Informatics, The University of Tokyo, JAPAN

Abstract

It is needed for humanoid to acquire not only just a trajectory but also aim of the behavior and symbolic information during behavior development. We have proposed the mimesis system as a framework of synchronous learning model for behavior acquisition and symbol emergence. However, the motion elements which are fundamental representation of behavior have stood on the unsuitable assumption that they are given without taking robots' embodiment and dynamics into consideration. In this paper, the design theory of motion elements with consideration of the embodiment are shown, and novel methods of realization of the mimesis for real humanoids is proposed.

1 Introduction

It is a major challenge of contemporary robotics to study robotic behaviors and intelligence in the full scale of complexity mutually sharing research outcomes and hypotheses with human behavioral science and human intelligence.

The discovery of mirror neurons[1] has been a notable topic of the brain science. The neurons in the brain of a primate fire when the other's particular behavior is observed as well as when the same behavior is performed by himself. The fact suggests that the behavioral cognition process and behavioral performance process are tightly coupled. It may even leads to the hypothesis that they appear as results of a single process of information processing. If they do, the information process, to the authors' point of view, would be a source of proto-symbol that is well grounded to the physical world and embodied using his own body.

It is also known that a newborn infant immediately starts communication with his mother by imitating her performance, and develops the skill of communication and performance. It suggests the relationship between the symbols for communication and the behavioral acquisition through imitation and communication. T. Deacon[2] pointed out that the language as communication using symbols and the brain are co-evolved. M. Donald[3] claims that symbol manipulation and communication using symbols are the origin of the mind.

The study of mechanism of information processing for behavioral cognition and behavioral performance may be an entrance to the synthesis of human-like intelligence and the brain-like information processing. The acquisition process using the mechanism has already been a field of robotics study known as imitation learning[4][5][6]. We would like to term the whole process of the information processing of behavioral cognition and performance, and the acquisition of behavioral performances and the associated symbols, as the principle of mimesis.

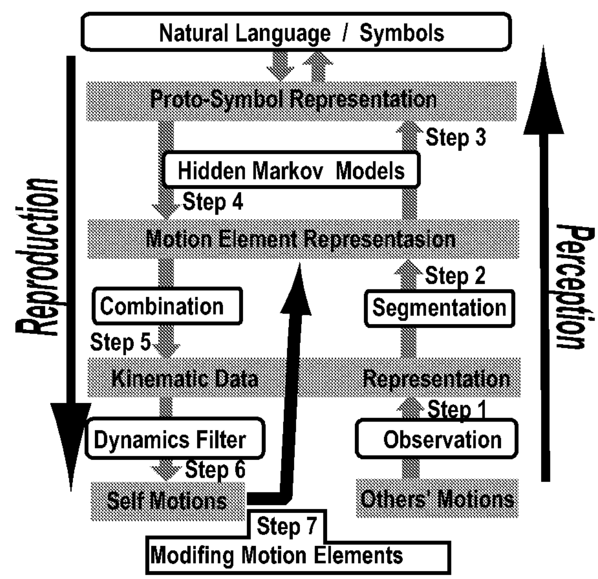

Figure 1: Overview of the mimesis system

We have proposed the mimesis system for realizing the function of mimesis in an engineering point of view[7][8]. This system is based on HMMs, recognizes and generates time series data, and emerges symbol expression called primitive symbol. However, in this system, motion elements for expressing whole body behavior was beforehand given by the developer, and they are not changed. In this paper, we flexibly reconstruct the motion element to be suitable for the kind of observed behavior and a robot's embodiment. In order to realize the upper things, we propose a new design theory through two techniques, introduction of the mathematical technique called mixture Gaussian HMMs, and evaluation of the achievement condition of the purpose of behavior.

2 Acquisition and Embodiment of Motion Elements

2.1 Outline of mimesis loop

We have propose a framework which can abstract whole body motion into certain symbols, generate motion patterns from the symbols, and recognize total behavior. The framework is referred to as mimesis-loop with the concept of the mimesis, as shown in Fig.1.

The mimesis system consists of two phases. In the first half, observed motions are transformed into sequence of motion elements (Step 1,2), and symbol representation is generated from the sequence (Step 3). We call the representation as proto-symbols because that can be regarded as primitive symbol representation comparing with tight symbol representation like languages. In the latter half, sequences of motion elements is decoded and reproduced from proto-symbols (Step 4). After that, each motion elements are integrated (Step 5) and transformed into natural motion representation with consideration of dynamics conditions (Step 6)[9].

We have proposed a mathematical framework for the integration of motion abstraction and motion embodiment using by Hidden Markov Models (HMMs)[10]. The tight connection between motion perception (abstraction) and motion generation (reproduction) can be regarded as a computational model of the mirror neurons.

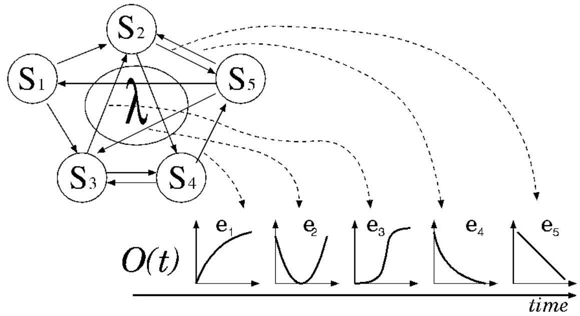

Figure 2: Motion Sequence and Hidden Markov Models

2.2 HMMs and proto-symbols

HMMs take time series data as input, and output probabilities that the data would be generated by the HMM. This probability is referred to as likelihood, which is key parameter for both recognition of others' behavior and generation of self-motion.

In this paper, we use two kinds of HMMs, discrete HMMs and continuous HMMs. First, discrete HMMs (DHMMs) are explained in order to introduce continuous HMMs (CHMMs). DHMMs have state transition model which cannot be observed. Each transition process output a certain symbol probabilistically, thus a sequence of symbols is output from the series of the transition as shown in Fig.2. The characteristics of DHMMs is represented by three parameters as follows:

1. $A = \{a_{ij}\}$ is a states transition matrix. a_{ij} indicates a probability of transition from state q_i to state q_j.

2. $B = \{b_{ij}\}$ is a output probability matrix. b_{ij} indicates a probability of output symbol o_j from state q_i.

3. π is a vector of probability for initial state distribution .

The HMMs, namely above three parameters, can represents the background dynamics of motion sequences

$$\mathbf{O} = \{o_{k(1)}, o_{k(2)}, \ldots, o_{k(T)}\} \tag{1}$$

by means of the combination of the each motion element e to each output symbol o, where T is time period of the motion sequence. Therefore, we have defined the proto-symbol P_S as follows.

$$P_s \stackrel{\text{def}}{=} \{A, B, \pi\} \tag{2}$$

The three parameters are learned based on the criterion that the likelihood value $P(\mathbf{O}|\{A,B,\pi\})$ ought to be the highest value where the $\mathbf{O}$ is the observed sequential symbols. DHMMs can calculate the probability that a sequential symbol $\mathbf{O}$ is generated by the set of $\{A,B,\pi\}$, which is referred to likelihood value $P(\mathbf{O}|\{A,B,\pi\})$, which is used in the motion recognition process.

2.3 Weak points of the previous system and methods of refine it

We had faced several problems on the proposed mimesis system [7][8] as followings:

1. There is no exact principle how to design the motion elements. In previous works, static and limited motion elements had been embedded by the developer beforehand. It is necessary for the mimesis system to develop the motion elements in order to be suitable for imitation learning from observation experience.
2. Physical condition of the motion had not been taken into consideration. As physical characteristic of learner and demonstrator is different, therefore, the observer cannot reproduce the same motion. The motion elements have to be suitable for both recognition of other's motion and embodiment of humanoids.
3. The motion elements were correspond to each joint. Thus the number of motion elements becomes no less than the number of DOFs in order to represent the whole body motion. it causes the complexity of symbol representation. Furthermore, the motion elements should represent correlation information between each joint.

Because of the above-mentioned reason, we propose a new method for acquisition of motion elements with two characteristics; a) use of continuous HMMs and b) modification of elements during observation and generation loop. The approach a) enables the system to avoid the problems 1) and 2). On the other hand the approach b) enables the system to avoid the problem 2) and 3).

To make motion elements with consideration of humanoid embodiment, we modify motion elements based on evaluations of reproduced humanoid's behavior. This process correspond to Step7 of Fig.1. By the repetition of these processes, motion elements are involved with consideration of humanoid embodiment. The detail of two methods are mentioned in Chapter 3 and 4 respectively.

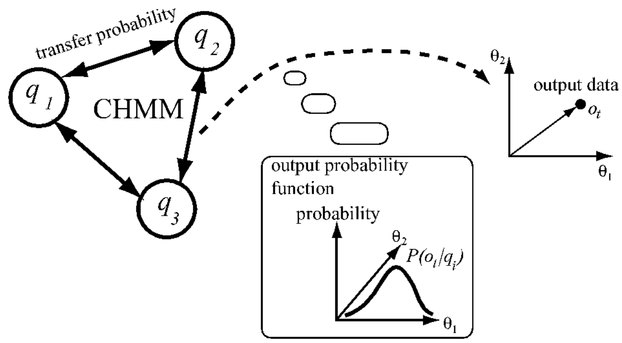

Figure 3: Outline of continuous Hidden Markov Models

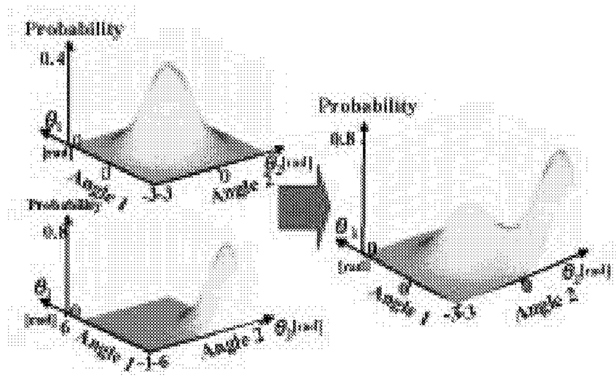

Figure 4: Concept of combination of motion elements. (In case that degree of freedom is two.)

3 Motion Elements Development using CHMMS

3.1 Continuous Mixture Hidden Markov Models (CHMMs)

We introduce continuous HMMs which is a kind of HMMs which can treat continuous data. The difference between DHMMs and CHMMs is that the transition process outputs continuous vectors as shown in Fig.3. Therefore, output probability matrix B becomes probability density functions in CHMMs. Here, the density function is approximated with linear combination of Gaussian functions in order to make the density function simple.

$$P_i(\mathbf{o}) = \sum_{j=1}^{m} c_{ij} G_{ij}(\Sigma, \mu) \tag{3}$$

Where $P_i(\mathbf{o})$ is probability density function for i-th state node to output the continuous vector o, $G(\Sigma, \mu, \mathbf{o})$ is the Gaussian function:

$$G(\mathbf{o}) = \frac{1}{\sqrt{(2\pi)^D |\Sigma_{ik}|}} \exp\{-\frac{1}{2}(\mathbf{o} - \mu_{ik})' \Sigma_{ik}^{-1} (\mathbf{o} - \mu_{ik})\} \tag{4}$$

Σ is the covariance matrix of continuous vector, μ is the mean of continuous vector, m is the number of mixture Gaussian functions, c_{ij} is mixture coefficient, and D is the number of dimension of continuous vector $\mathbf{o}$.

The only different point from DHMMs is that the learning target changes to $\{\pi, A, c, \Sigma, \mu\}$. These parameters are calculated using Baum-Welch algorithm which is one of EM-algorithm.

3.2 Applying CHMMs to mimesis system

Here, we divide the parameters, then redefine the proto-symbols and motion elements as follows.

$$P_s \stackrel{\text{def}}{=} \{A, \pi, c\} \tag{5}$$

$$e_i \stackrel{\text{def}}{=} \{\Sigma_i, \mu_i\} \tag{6}$$

In other words, kinematic representation of the motion consist of basic motion elements whose number of kind is as same as the number of mixture Gaussian component. An important issue is that the motion elements are automatically calculated by the Baum-Welch algorithm. Additional to that, the relation and correlation among each joint can be represented by the motion elements because each elements hold information of whole body motion and correlation matrix Σ. In previous works, the number of motion elements m is fixed, and entrusted to the developer, however, CHMMs can also optimize the m.

A concept of combination of motion elements is shown as the Fig.4. Fig.4.1 shows motion elements as multivariate Gaussian functions, and Fig.4.2 shows probability density of output motion vector. In previous works of mimesis loop, we prepare sets of motion elements beforehand. And we abstract sequences of motion elements with DHMMs. To sum up, mimesis system can take following advantages with CHMMs.

- Various motion patterns can be made by combination of continuous motion elements.
- A motion element is able to express the whole body motion whose degree is as same as the degree of target motion.
- Pattern of motion elements and the number of motion elements is automatically calculated.

4 Closing the mimesis loop for embodiment

We integrate the behavior perception and behavior generation. In other words, the integration is corresponds to closing the mimesis loop. It is needed for new born babies to continue trial and errors in order to acquire self behavior. In a similar approach, the humanoid can acquire inner parameter of mimesis system with the consideration of embodiment through repetition of behavior perception and generation. We prepare three criteria for evaluation of generated behavior as followings;

1. Whether the generated trajectory is resemble to the target motion, or not.
2. Whether the performance of the recognition and generation process increase, or not.

We use evaluation criterion for first condition as follows:

$$E_\theta = \int |\theta_{in}(t) - \theta_{out}(t)| dt \tag{7}$$

Where $\theta_{in}(t)$ and $\theta_{out}(t)$ indicate the joint angle of an observed ideal motion and a generated motion. For second condition, we adopted following criterion:

$$E_R = R^{i+1} - R^i \tag{8}$$

where R^i indicates recognition ratio in i-th learning loop. When the sign of E_R is plus, the mimesis system judges that i-th motion data is suitable for recognition process, adds the motion data into database, and calculates the motion elements again.

A method for reproduction of the behavior is as same as the previous works. See the details of the method in bibliography [7][8].

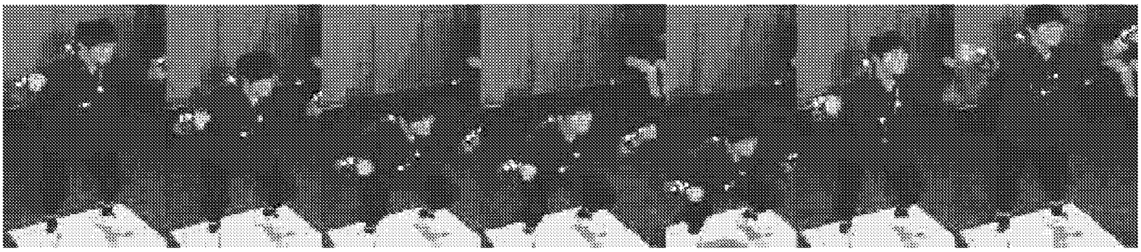

Figure 5: Motion Capturing System: squat motion for learning data

5 Simulation results

We have confirmed the performance of our method. 20 squat motions are observed as learning sample data using by the *Behavior Capturing System* as shown in Fig.5. The time period of each motion is about 2[sec] with sampling time 20[msec]. In this experiment, we put the motion elements in charge of three joint; hip(pitch axis), knee and ankle(pitch axis), that is, the dimension of the motion elements is three. We have set the number of state nodes in the CHMM ten, and the number of mixture for each node ten, that is, 100 motion elements should be created.

After the 20 observations, motion generation process is executed 80 times, and appropriate motions are added into database. Figure 6 shows the acquired motion elements after the 80 times generation. Each dot mark indicate the motion element, solid line indicate the original motion's trajectory. As the figure shows, the motion elements are located near the original motion, that is, our method shows good performance.

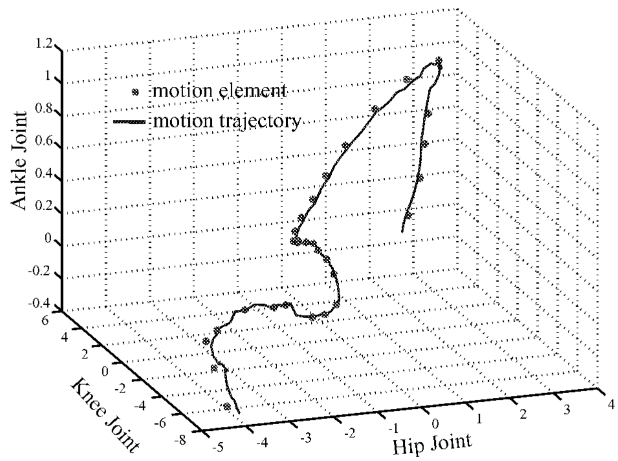

Figure 6: Motion elements and original motion

Using the method, we also performed the experiment with walking behavior as shown in Fig7. Fig.8 shows experimental result on a humanoid robot. The generated behavior is suitable for the real robot because the acquisition process considers the embodiment of the humanoid in dynamics simulator. The experiments are executed with its body dangled, because any balance controller are not considered in current stage.

6 Conclusions

In this paper, two improvement was performed for the refinement the mimesis system. One is automatical generation of motion elements using continuous HMMs. Another is that the system gained the embodiment through modification of motion elements based on the evaluation of the generation motion. We confirmed that out approach aims at suitable direction through two basic experiment.

There is vector quantization (VQ)[11] as a standard strategy for the conversion from continuous data into discrete data. VQ is convenient method, however, there is no guarantee that the generated discrete data are suitable for the embodiment of the humanoid. the method proposed in this paper considers the both recognition process and reproduction process, thus the acquired motion elements are suitable for the humanoid. In addition to that, the proposed method has another advantage that the number of motion elements is optimized using by the continuous HMMs.

Figure 7: A observed motion of walking

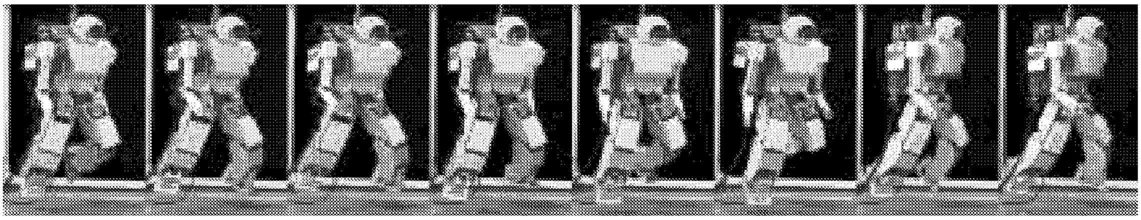

Figure 8: A generation result of walking motion on a humanoid robot

Acknowledgement

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

References

[1] V. Gallese and A. Goldman. Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2(12):493–501, 1998.

[2] T.W. Deacon. *The symbolic species.* W.W. Norton & Company. Inc., 1997.

[3] M. Donald. *Origins of the Modern Mind.* Harvard University Press, Cambridge, 1991.

[4] S. Schaal. Is imitation learning the way to humanoid robots? *Trends in Cognitive Sciences*, 3(6):233–242, 1999.

[5] M.J. Mataric. Getting humanoids to move and imitate. *IEEE Intelligent Systems*, pages 18–24, 2000.

[6] Y. Kuniyoshi, M. Inaba, and H. Inoue. Learning by Watching: Extracting Reusable Task Knowledge from Visual Observation of Human Performance. *IEEE Transaction on Robotics and Automation*, 10(6):799–822, 1994.

[7] T. Inamura, Y. Nakamura, H. Ezaki, and I. Toshima. Imitation and primitive symbol acquisition of humanoids by the integrated mimesis loop. In *the Proc. of Int'l Conf. on Robotics & Automation*, pages 4208–4213, 2001.

[8] T. Inamura, Y. Nakamura, and I. Toshima and H. Ezaki. Mimesis embodiment and proto-symbol acquisition for humanoids. In *the Proc. of Int'l Conf. on Advanced Intelligent Mechatronics*, pages 159–164, 2001.

[9] K. Yamane and Y. Namamura. Dynamics filter – concept and implementation of online motion generator for human figures –. In *Int'l Conf. on Robotics and Automation*, pages 688–695, 2000.

[10] L.R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. In *Proc. of IEEE*, volume 77, pages 257–286, 1989.

[11] S.Young et al. *The HTK Book.* Microsoft Corporation, 2000.