Constructive Research of Active Perception by Cognitive Experiment and Simulation Using Neural Networks
(ニューラルネットワークを用いた認知実験とシミュレーションによるアクティブパーセプションの構成論的研究)

Department of General Systems Studies,
Graduate School of Arts and Sciences, The University of Tokyo
東京大学大学院総合文化研究科広域科学専攻広域システム科学系

OGAI, Yuta
大海 悠太

January, 2011

Contents

Chapter 1

Introduction

1.1 Ecological world view

In the book *"Encountering The World"* (Reed, 1996), Reed considered extensions of ecological psychology started by J. J. Gibson (Gibson, 1979). Reed criticized the traditional psychology based on the mechanistic world view, which attempts to explain life using mechanical metaphors. Reed wrote the mechanical systems given as the metaphors are telephone exchanges, steering servomechanisms, and digital computers "do not act unless put into action by an external agency" (Reed, 1996, p.9). Reed insisted that traditional psychologists neglect the problem of autonomous agency; he feels that mechanical systems are mere tools, and cannot move autonomously like animals and humans.

Going against traditional psychology, Reed suggested "ecological alternative based on the biological concepts of the regulation of activity" (Reed, 1996, p.9). Reed considered that autonomous agency as an ecological alternative does not have any central processes for constructing complex behaviors, and is realized by the self-regulation of activities afforded in environment. We call this viewpoint as "the ecological world view". Reed stresses the importance of the environment to understand intelligence, and gave the example of Darwin's research on earth worms to explain the ecological world view (Darwin, 1881).

In Darwin's research, worms showed unexpected intelligence with regard to various environmental changes, although they do not have a brain. For example, worms filled their nests in by pulling leaves from outside. As a result, air flow into the nests is prevented. Darwin assumed that this behavior was an adaptation to prevent the skin of worms from becoming dry in the nests. Moreover, worms pulled leaves by grasping their tips, but if the bases are narrower than their tips, the leaves are pulled by grasping the bases. If worms pull leaves by grasping their narrower edges, they can fill the nests efficiently.

From the observation of this adaptive behavior, Darwin insisted that although worms are simple creatures, they possess "some degree of intelligence" (Darwin, 1881, p.98). Reed put forward the ecological world view as more suitable than the mechanistic world view for understanding such animal behaviors because Darwin's worms realize their intelligence by perceiving and acting on the environment. We consider that action and perception are inseparably related in such animal behaviors. In this thesis, we discuss active perception from the ecological world view.

1.2 Active touch

We perceive by what we do, and we act by what we perceive. Repeats of this perception-action loop compose a whole perception. For example, we observe active perception in almost all modalities. We perceive shapes and texture by actively touching objects. We cannot stop small involuntary eye movements, when we see the world. We taste food by moving our teeth and tongue. The sense of smell strongly relates to breath. We can actively listen to what we want to pay attention to (e.g., the cocktail party effect (Cherry, 1953)).

We consider active perception in the sense of touch, which is more primitive than the other senses such as vision or hearing. In the book *"Der Aufbau der Tastwelt"* (Katz, 1925), Katz discussed tactile sensation, and analyzed active movements to perceive objects through various experiments of active touch. For example, Katz investigated finger movements when a subject was touching soot-covered papers(Figure 1.1) Katz noticed that all subjects left completely untouched places here and there between the touched places on the papers. Normally, when people touch paper, they cannot touch every part. Katz insisted that people can fill in those to have a coherent image.

After his research, psychological experiments of active touch have been carried out on by many researchers. For example, Gibson showed that when people use blind touching for cookie cutters, they can recognize the shape of them better if they move their hand actively. Therefore, Gibson claimed the importance of active touch (Gibson, 1962). Lederman and Klatzkey reported relations between features of objects (e.g., texture, weight, and shape) and human's hand movements to recognize the features (Lederman and Klatzky, 1987). Moreover, recent technology enables new psychological experiments. By using haptic interface (Robles-De-La-Torre and Hayward, 2001), Robles-De-La-Torre claimed that humans use force cues more than geometry to detect a bump and hole.

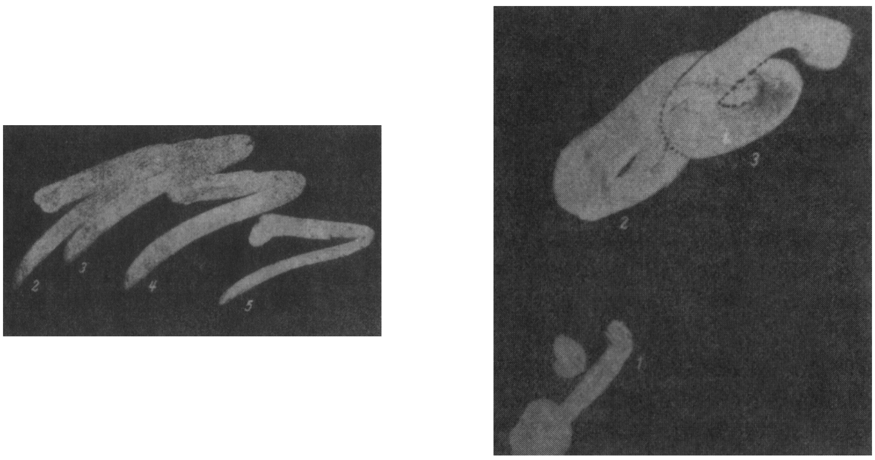


Figure 1.1: Finger Movements on soot-covered papers (from (Katz, 1925))

1.3 Four classes of active touch

Active touch has been intensively studied; however, there are some aspects in the active touch concept that are confusing. Therefore, we first categorize the concept of active touch into four classes in order to show what makes active perception possible. The classes are bound up with each other, but they can be considered separately by using proper methods of observations.

1) Static physical features

Active touch is controlled by static physical features like the shape of hands, size of objects, and the high resolution area in fingers. For example, in cutaneous mechanoreceptors, the densities of Meissner corpuscles and Merkel cells are high on fingertips (Johansson and Vallo, 1979). The effective utilization of small areas is one of the reasons that people touch objects by moving the fingertips. Moreover, in the research of Lederman and Klatzkey, when people feel an object by graving it, their perception depends on both the size and shape of the hands.

2) Temporal aspects of sensory inputs

To experience tactile sensation, people need to move hands freely in most cases. Then, cutaneous mechanoreceptors on the fingers receive time varying stimulation. In addition, people receive time varying stimulation as somatic sensation, when their hands are moved even by someone else.

In contrast to the research of active touch using cookie cutters by Gibson, some studies confirmed that people can detect objects by passive touch better than active touch. Schwartz et al. reported their subjects could also detect the shapes of cookie cutters presented passively on their hands (Schwartz et al., 1975). Moreover, Magee and Kennedy did experiments on the recognition of raised-line drawings (writing materials for blind people) to compare active and passive conditions, and have shown guidance is helpful and leads to better identification than unguided exploration (Magee and Kennedy, 1980). In the case of Schwartz et al., the subjects utilize time series data on the mechanoreceptors of their hands to process the sensory information. In the case of Magee and Kennedy, time series data on the mechanoreceptors of their fingers and somatic sensation of their arms are used. Thus, passive time series data is sometimes enough to detect objects. However, in daily life, objects normally stop, and people have to actively move their hands to receive tactile sensation of the objects. Therefore, an action to make time series data also means active touch.

3) Input and output relationship

For example, Turvey investigated people's perception of rod length by wielding it, and claimed people perceive its inertia tensor as an invariant (Turvey, 1996). In other words, the quantity is a tension between movements of the arm's muscles and resistance against the movements. The research of Robles-De-La-Torre et al. also applies the same idea in touches of a bump and hole (Robles-De-La-Torre and Hayward, 2001). These relationships are perceived only by self-generated movements.

However, we suppose that relationships between inputs and outputs do not always need to be simple invariants such as the inertia tensor, which appeared in the research of Turvey. We consider that people can also perceive more complex relationships than such invariants.

4) Higher order cognition

Active touch is also controlled by higher order cognition like prediction for self-produced stimulation and selective attention. For example, most people cannot tickle themselves. People feel a different sensation between self-produced and externally produced stimulation. In the case of self-produced stimulation, people predict the stimulation, and suppress it. Blakemore et al. used fMRI to examine the neural responses when subjects experienced a tactile stimulus that was either self-produced or externally produced, and found less activity in the somatosensory cortices and cerebellums when the stimulus was self-produced (Blakemore et al., 1998). In addition,we think phenomenon such as selective attention in the cocktail party effect (Arons, 1992) is also involved.

1.4 Dynamical system

A dynamical system has a fixed rule that signifies a future state from the current state, and is useful to understand active touch because it can describe a time series and relation between inputs and outputs in the classes of active touch. Its behaviors in the low-dimension are classified as chaos, limit cycles, quasi-periodicity, and fixed points. In high-dimensional dynamical systems, itinerant motion among varieties of the low-dimensional ordered states through high-dimensional chaos is commonly observed, and is known as chaotic itinerancy (Kaneko and Tsuda, 2003). Agents in the dynamical system acquire cognitive behavior which respond well by the features of dynamical system. In this thesis, we adopt recurrent neural networks (RNN) as the candidate for modeling cognitive behavior. RNN is a class of dynamical systems that can emulate most low and high generic behaviors of dynamical systems. Further, a dynamical system is convenient for real-time operation and an open environment.

A dynamical systems approach has already been applied in developmental psychology. Thelen et al. observed young infant reaching movements, and claimed the behavior emerges from the individual intrinsic dynamics of each infant (Thelen et al., 1993). Taga et al. studied spontaneous movements of young infants who have not yet acquired voluntary movements, and found that the general movements have chaotic dynamics by analyzing their time series (Taga et al., 1999). We also examine active touch as a dynamical system in this thesis.

1.5 Constructive approach

According to Hashimoto et al., a constructive approach is "a scientific methodology in which an objective system is to be understood by constructing the system and operating it" (Hashimoto et al., 2008, p.111). We apply the constructive approach to understand active touch.

G. Walter made "Turtle" robots by designing electric circuits, and demonstrated the importance of interaction between the bodies of the robots and the environments (Walter, 1950). Since this cybernetic era, the constructive approach has attracted attention in various fields. When R. Brooks proposed the subsumption architecture (Brooks, 1986), this tendency was stressed. The subsumption architecture is a way to decompose intelligent behavior into basic modules of simple reflective patterns, and dispose of the modules as layers to construct various kinds of macro behaviors that are not just reflective.

Robots designed using the subsumption architecture can move nicely by synergetics of the layers. Making robots with the subsumption architecture, Brooks insisted that robots do not need representations to have intelligence (Brooks, 1991b), and importance of embodiment (Brooks, 1991a). The robots receive immediate feedback as a result of their actions, so they do not need to have the whole model of the world. His research indicates the usability of the constructive approach as opposite to the conventional artificial intelligence. The idea of Brooks fits to the ecological world view. Clark also noted similarities between the insistence of Brooks and the ecological psychology of Gibson (Clark, 1997).

After Brooks's research, researchers tried to construct robots using the bottom-up approach in the same line. Sensory motor coordination (SMC) is one of these methods of robot design. Behavior of a robot on SMC is directly guided by the sensory input (Pfeifer and Scheier, 1999). SMC is compatible with the learning of a primitive connection between sensory as input and motor as output. The connections are generally composed of learning machines such as a neural network. SMC controlled by a neural network can show more sophisticated behaviors compared with classical systems composed using simple feed-forward and feed-back loops, and the behaviors become unpredictable due to features of the dynamical system of the neural network. Moreover, behaviors of such embodied agents and neural networks can be represented by a dynamical system.

To consider whether an agent has intelligence, it is important to observe the agent's flexible responses to externals, which include various environments and movements of other agents. If what an agent tries to learn is too easy for him, it becomes a mere optimization problem, and only predictable results are produced. On the contrary, if the task is difficult, the agent needs to have enough complexity to acquire abilities to complete it. The complexity of the agent is represented as the size of neural networks and different structures of bodies. For example, simple neural networks with only one hidden neuron cannot solve difficult tasks, but neural networks with many hidden neurons may solve difficult tasks, although it becomes hard to calculate the networks. Therefore, the difficulty of the tasks and the complexity of neural networks must be designed carefully.

1.6 Construction of ecological world

Varela et al. discussed the circularity between action and perception in terms of an observer and object, which is similar to an enactive approach (Varela et al., 1991). The subsumption architecture proposed by Brooks can also be taken as the enactive approach according to Verela et al. We consider that the enactive approach deals with the construction of the ecological world, which has circularity between the agents and environment. In this thesis, we construct the ecological world not by the subsumption architecture, but by modeling with

SMC and RNN.

It is also possible to apply the constructive approach to psychological experiments by constructing environments around subjects. In this thesis, we took the constructive approach for both modeling and psychological experiments.

1.6.1 Modeling

In modeling active touch, we combine neural networks, physical body, and environments to create a system without putting any intention or other secret ingredients. Most studies on active perception by modeling deal with active vision and touch. Therefore, we will concentrate on these areas. Differences of active vision and touch in modeling is mainly due to whether an agent can sense far objects. Therefore, we can apply the same modeling strategy to both problems. For example, Kato and Floreano investigated active vision using an evolutionary simulation model in which a neural network moves view and scales it, and discriminates between a triangle and a square (Kato and Floreano, 2001). Marocco and Floreano also investigated active vision using an evolutionary robot approach (Marocco and Floreano, 2002). Morimoto and Ikegami discussed dynamical categorization, which is as a self-centered categorization achieved by developing adequate sensori-motor couplings (Morimoto and Ikegami, 2006). The agents consist of homogeneous elements mutually connected by springs, and autonomously move by actively using these elements. In addition, Ikegami simulated an agent that has the Fitz-Hugh-Nagumo neuron network and moves around by receiving patterns on a floor as inputs (Ikegami, 2007). He discussed a relationship between active perception and "embodied chaotic itinerancy", which is named after its spontaneous selection of motion styles.

In Chapter 2, we use the modeling approach for microslips, which are psychological phenomena in which fluctuation occurs inevitably. Microslips are frequently observable phenomena. They were first named and observed in detail by Reed et al. (Reed et al., 1993). We simulate microslips as the computation model, and reinterpret Reed's observations in terms of dynamical systems.

1.6.2 Psychological experiments

The constructive approach in psychology is to make a virtual experience by constructing artificial systems. Robles-De-La-Torre dealt with the haptic interface to make a paradoxical situation of a bump and a hole (Robles-De-La-Torre and Hayward, 2001). The constructive approach in psychology has the advantage of generating various perceptions on the special or paradoxical situation through the man-machine interface. In Chapter 3, to investigate active touch, we use evolutionary computation to create the feelings of tactile textures, which are represented as Japanese onomatopoeias, and control parameters of the system which generates the feelings.

Chapter 2

Microslip as a Simulated Artificial Mind

2.1 Slippery action

Behavior is organized by action primitives. For example, making a cup of coffee consists of stirring the coffee, filtering the water, picking up the coffee cup, etc. Those primitives are not linearly linked but can be composed in more complex ways. We often experience an unexpected action selection/production that is different from our intention. This slippery action is what we know as a "microslip" (Reed et al., 1993). A slippery action or an action stutter in general, such as a slippery word, has been noticed in behavioral psychology (e.g., (Norman, 1981; Reason, 1989)), and was named a microslip afterwards by people doing "ecological psychology," which was started and developed by Gibson (Gibson, 1966), and Reed (Reed, 1996).

In Reed et al.'s example of making a cup of coffee, microslips are observed when a man touches a cup, then instantly detaches from it, then touches it again, or when a man reaches for a spoon but does not touch it and instead touches the coffee powder. These are examples of microslips frequently observed in making a cup of coffee. It is not a "macro" action pattern but a micro movement that emerges about once per a minute in a normal condition. Microslips occur in everyday life and are phenomenologically classified into 4 patterns by Reed et al.(Reed et al., 1993). For example, making a cup of coffee includes an action pattern such as "grasping a coffee cup." The following microslips are associated with this grasping action:

hesitation
: a faltering attempt to picking up a cup just before making contact.

trajectory shifts
: readjusting the direction toward the target cup.

action stutters
: withdrawing from the target or making contact with a different object.

hand-shape changes
: the hand shape changes while reaching for the target.

More detailed classification can be found in Reed et al.'s original paper. However, the question arises naturally whether those classifications are inevitable and, if so, to what degree. Suzuki (Suzuki, 2001) re-classified microslips and unified the "subtle withdrawing" cases (hesitation and action stutters) and "in-motion" cases (trajectory shifts and hand shape changes) . He further classified actions into upper (a global unit of actions) and lower tasks (a local unit that composes the global unit), and argues that the microslips of withdrawing cases are frequently observed in-between lower tasks. This is consistent with the view that tasks are hierarchically organized. However, such hierarchy organization is not a static structure but changes from time to time, which we will focus on in this paper using the dynamical systems approach.

We see the mechanism of microslips as a fluctuation of action production or a breakage of recursive action synthesis, because we think that a microslip implies a human's difficulty in doing the same action recursively. Robots or machine arms do not usually have microslips. Induction of slippery action may be the fundamental aspect of human cognition. In other words, action production is always associated with fluctuation, but this fluctuation is a part of the normal state of action production. We think a microslip is not an error action and argue that human action production should be understood as a more dynamic and open-ended process.

In the next section, we describe the computational model of microslips and simulate it to search for new perspectives. As a result, we propose a new characterization for a microslip. The following is a summary of how we interpret microslips in our formulation.

- A microslip is characterized by an entropy measuring the mixture of two intentionalities.
- A microslip is a reflection of the cognition sensitive to the spatial layout of objects.
- A microslip is organized by a heterarchy (i.e. more like a network organization; Jen (Jen, 2003)), not a hierarchy, of actions.

Here we model microslips as a dynamical system. In general, understanding cognitive behavior in terms of a dynamical system has advantages and disadvantages. The disadvantages come because dynamical systems modeling becomes possible only by reducing the abundant complexity of a given cognitive phenomenon, which may miss essential points. For example, even making a coffee is too complex a task to be modeled as it is. The advantages are that we can analyze the behavioral pattern in terms of the established concepts, such as entropy and fractal basin boundaries. In this paper, we simulate an environment with only two objects, and an agent tries to pick up one of the two objects. Even though the task is simple, the agent shows interesting behaviors, which we think refer to the inherent complexity underlying microslips. The concepts of dynamical systems modeling have been used to understand generic cognitive behaviors (e.g., (Port and Van Gelder, 1995)), in particular as a conceptual tool to study developmental processes by Thelen & Smith (Thelen and Smith, 1994). Especially, a concept of chaos and chaotic itinerancy has been explored in brain science (e.g., (Tsuda, 2001; Freeman, 2001)) and mutually interacting agents (e.g., (Ikegami and Morimoto, 2003; Iizuka and Ikegami, 2004)). In the field of cognitive robotics, the dynamical systems concept has been widely accepted (Pfeifer & Scheier (Pfeifer and Scheier, 1999)), in particular by Tani & Fukumura (Tani and Fukumura, 1997) for the first time to show the existence of chaos in a robot experiment. This paper is on the track of understanding cognitive behaviors in terms of dynamical systems modeling, in particular as a singular property of a certain type of dynamical systems.

After introducing modeling in section 2.2 and the results and analysis in section 2.3, we discuss the concept of singular dynamics and hierarchy in section 4 to understand the dynamic perspectives of the behavior and mind.

2.2 Modeling

2.2.1 Task and environment

In this section, we introduce a dynamical systems modeling of microslips. The mathematical framework we study here consists of an agent going to pick up objects in a two-dimensional field. An agent has a simple neural network and determines which objects to take by detecting where the objects are. Since microslips are observed in a rich environment (e.g., with more than two objects in an environment), we study the case of two objects as a minimal example. An agent has an "intention" to pick object 0 or 1 and chooses an action primitive associated with the intention. As we discuss below, an agent has two neural nodes (S_0 and S_1), and the agent goes for object 0, if the value of state S_0 is larger than that of S_1. We thus call the state of the agent the "intentionality for getting the object 0" and the associated action pattern "action primitives", since an action pattern segment does not have a meaning, but as a whole, it constitutes an action sequence of a given intentionality.

The spatial trail of an agent reaching an object (0 or 1) is illustrated in Figure 2.1. An agent has to constantly receive a distance of an object (or objects) to synthesize the action pattern. This is computed by a discrete time recurrent neural network (see the details in section 2.2.2) and the network evolves using a standard genetic algorithm (GA) (see the detail in section 2.2.3). To evolve the network, we adopt three tasks for the same agent to accomplish.

Task 0
: A single object 0 is on the field, and the agent has to get it.

Task 1
: A single object 1 is on the field, and the agent has to get it.

Task 2
: Two objects are on the field, and the agent has to get one of the two.

With tasks 0 and 1, the agent learns an action primitive for each of the objects, and the agent is expected to combine the acquired action primitives to achieve task 2. Since task 2 requires the agent to get either one of the objects, the agent is free to choose without any biases. If an agent can evaluate the necessary time to reach the object, the agent merely tries to take the closest object without any hesitation in task 2. However, an agent sometimes does something different, which we will focus on in this paper.

2.2.2 Recurrent Neural Network

We used a standard discrete time recurrent neural network (RNN) with 3 layers in this experiment (Figure 2.2) to design an agent's internal dynamics. The input layer (the first layer) gets the x- and y-coordinates of the two objects in the field [1] and two recurrent states. The neural states on the 2nd and 3rd (output) layer $u_{i,j}, (i = 2, 3)$ are determined by the following equation (2.1):

$$u_{i,j} = g(\sum_{k=1}^{N_{i-1}} w_{i-1,j,k} u_{i-1,k}) \tag{2.1}$$

where $w_{i-1,j,k}$ is the connection weight between $u_{i-1,k}$ and $u_{i,j}$ and N_{i-1} is the number of neurons on the $(i-1)$th layer. The function $g(x)$ is a sigmoid function given by the following equation (2.2):

$$g(x) = \frac{1}{1 + e^{-\beta x}} \tag{2.2}$$

Perceiving the objects in the environment, an agent moves forward (see equation(2.3)(2.4)) by updating the force vector, (F_{0x}, F_{0y}) and (F_{1x}, F_{1y}) every discrete time step. Two choice states S_0 and S_1 are also updated to determine which force vector to use. The amplitude of the forth vector and the choice states are given by the assigned neural states of the output layer (i.e. the neurons on the third layer $i_{3,j}$).

When there is one object in the field, the other inputs are suppressed. For example, in case of task 0, the x- and y-coordinates of object 0 are given as (O_{0x}, O_{0y}) and (O_{1x}, O_{1y}) have constant null entries [2] and only (F_{0x}, F_{0y}), are used for navigation. But in task 2, the choice states are crucial, because the agent has to continuously select either object 0 or 1, which is determined by the choice states. If $(s(0) > s(1))$, then (F_{0x}, F_{0y}) is chosen; otherwise, (F_{1x}, F_{1y}) is chosen. This choice procedure is executed at any time, so that the agent temporally switches between (F_{0x}, F_{0y}) and (F_{1x}, F_{1y}). When an agent finally comes to the neighborhood (*delta*) of an object within a given time range, we say that "the agent has reached the object."

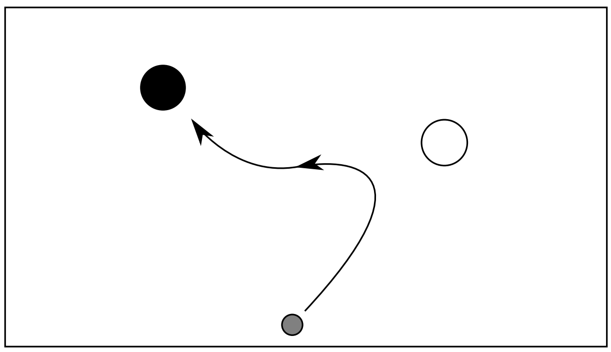


Figure 2.1: Simulation Environment. An agent moves in this environment to get an object (i.e. a black and white circle in the figure).

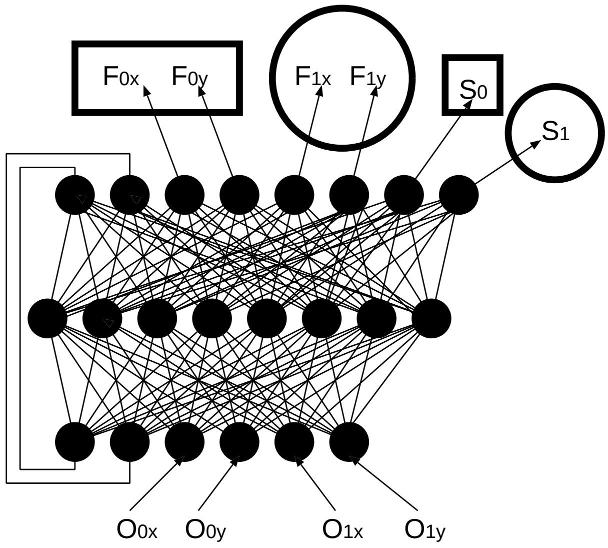


Figure 2.2: An illustration of RNN used in this simulation.

The agent has a velocity vector (V_x, V_y) and a virtual mass of m, and the equation of motion of the agent is given by the following:

$$x^{'} = x + V_x \Delta t \qquad V_x^{'} = V_x + \frac{F_x \Delta t}{m} \tag{2.3}$$

$$y^{'} = y + V_y \Delta t \qquad V_y^{'} = V_y + \frac{F_y \Delta t}{m} \tag{2.4}$$

where (F_x, F_y) is either (F_{0x}, F_{0y}) or (F_{1x}, F_{1y}).

2.2.3 A genetic algorithm

We used a standard genetic algorithm (GA), which is known as a core tool for evolutionary robotics (Nolfi and Floreano, 2000). Taking each neural weight as a gene, we selected the best 4 individuals out of 20 to breed the next population per each GA generation. No cross-over is adopted here; only a point mutation (of the normal distribution) is adopted. The fitness function $G(t)$ is given by the following equation (2.5) for each task t.

$$G(t) = \sum_{i=1}^{N} (T_i(t) + D_i(t) + P_i(t)) \tag{2.5}$$

where N is the total number of trials for each task by randomly assigning objects in the field (we used $N = 10$). For each trial i, $T_i(t)$ ($< Tmax$) gives the amount of time to get the object, and $D_i(t)$ gives the distance from the object at the maximum time limit ($Tmax$). $P_i(t)$ is a punishment if the agent can not get the target object (i.e., $(P_i(t) >> T_i(t), D_i(t))$). The total fitness G is a sum of the function of each task (i.e., $G = G(0) + G(1) + G(2)$). The lower the value of G, the better the agent performs.

2.2.4 A definition of simulated microslips

An ideal solution of the agent's behavior is expected to get the closest target. However, the spatially closest object does not mean that it requires the shortest time. Furthermore, the required time to get to an object is different when there are two objects. Therefore, an agent's choice behavior becomes complicated.

When an agent takes more time to get a target object, we interpret that the agent is in a hesitating state (or an agent has more fluctuation of action selection). In this model, we define microslip as this fluctuation of the action selection process. That is, we identify microslips when the following are observed.

- Frequent switching between two action primitives.
- Complex reaching style (agent's navigation pattern) to get an object.

As we will see, microslip is a function of the spatial "layout" of objects. A subtle difference in the layout determines the degree and styles of microslips. Since the agent constantly takes in the relative distance of the objects as inputs, we say that the relative layout of objects affords the agent's preference. Thus, our second message from this study is that affordance of layouts (Fukuma, 2003; Reed, 1996) sensitively controls microslips.

2.3 Results

2.3.1 An evolution of RNN by a GA

Figure 2.3 describes an example of GA trials, an evolution of the best agent and the average number of action switching events while processing each task. The number of action switching events is computed by the frequency of alternating between the $So > S1$ state and the $S1 > S0$ state averaged in time.

An enlargement of Figure 2.3 is given in Figure 2.4, where we notice that the fitness improves in the order of task 1, task 0, and, lastly, task 2, which is intuitively a correct order. After the 30000 GA generation, objects in almost any position become accessible by the agent. It should be remarked that sometimes an agent fails to get objects, which is implied by the abrupt increases of fitness values in Figure 2.4. These failing behaviors can be observed around 50000-90000 generations and also around 120000-200000 generations. Usually, agents in those generations have microslips. In the following sections, we pick up the agent from the 160000th GA generation and examine its behavior in detail.

2.3.2 Motion patterns in task 2

We give variations of behaviors in task 2 (two objects case) in Figure 2.5, which needs careful classification of the behavior. The initial position of the agent is fixed, and the relative positions of the objects are used as x- and y-coordinates in the figure. In particular, when we fix the y-coordinates of the objects at 225 and vary the x-coordinates of the two objects from -300 to 300, we can examine the agent's behavior in the 2-dimensional plane (the x-coordinate of an object 0 as the horizontal line and the y-coordinate of the object 1 as the vertical line).

We use different line types for the different action primitives against object 0 (solid line) and object 1 (broken line), where object 0 is black and object 1 is

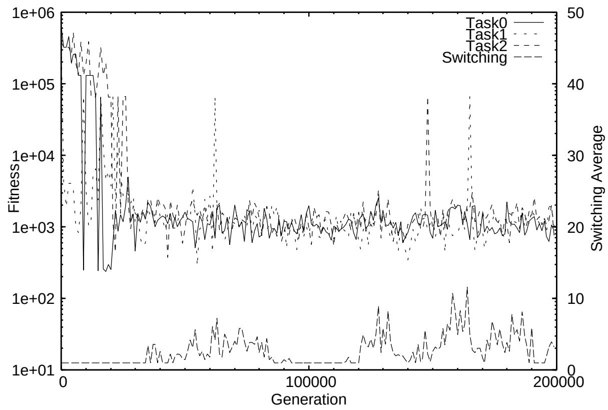


Figure 2.3: Evolution of the fitness value of each task and the number of action switching events in average (the lower line).

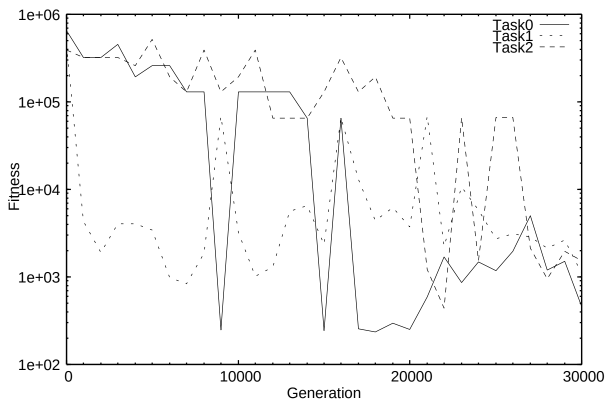


Figure 2.4: Evolution of the fitness value of each task.

white. We will call the relative position of the two objects the spatial "layout" of the objects hereafter. There are two different styles of behaviors in task 2: straight approaching cases (Figure 2.5 a) and b)) and complex action switching cases (Figure 2.5 c) and d)). We define the latter cases as "microslips" by our definition. The emergence of microslips depends on the spatial layout of the two objects, which we study in detail in below. The first analysis is a time distribution of reaching objects (Figure 2.6).

Figure 2.6 is a diagram of the number of time steps required in task 2 (to get the either of the objects) in the 2-dimensional plane organized by each x-coordinate of the objects. Namely, those points in the diagram correspond to the layouts of the objects with a common y-coordinate at 225. The spatial layout of the objects with the darker color requires more time to reach either of the objects. In the case of tasks 0 and 1, the required time steps are also represented in this figure. Above, this figure is provided a time required in task 0, and the left-hand side of this Figure 2.6 is provided the time required in task 1. It is interesting to note that Figure 2.6 of task 2 is not a simple superposition of tasks 0 and 1.

If an agent can compute the closest object from the inputs, Figure 2.6 should be a superposition of tasks 0 and 1. But an agent seems to compute something different so that the reaching behavior becomes a complex function of the layouts. To clarify this point, we compare the required time to get an object 0 in tasks 0 and 2 in Figure 2.7. We see that the required time for reaching object 0 is drastically perturbed by object 1 in task 2. This interference between objects 0 and 1 is the source of this complex basin structure. You may think that task 2 is solvable by simply neglecting the other object in this region. But the rest of the region shows a complex basin structure, which suggests a strong interference between inputs from two objects, if they exist simultaneously.

It should be noted that the fitness requires the shortest time to get an object, but the mapping between the time and the distance for an agent to get an object does not form a simple function. Therefore, sometimes a complex selection pattern has evolved. But this is not always the case. More sluggish selection patterns can appear, such as "always taking object 0 in task 2". In other similar experiments (e.g., (Cangelosi et al., 1994)), the conflict between the two possible selections can be resolved by attaching priorities or having an adequate selective attention. We have not introduced any priorities here, and thus the situation is more difficult. In the Appendix A, we give other selection patterns from different generations and in the Appendix B, we also give other pattern obtained from the different GA runs.

Figure 2.8 represents the final outcome of task 2. The black region in Figure 2.8 corresponds to object 0, and the white region corresponds to object 1. The gray region corresponds to the unselected region (none of the objects were selected). Comparing Figures 2.6 and 2.8, we notice that the layout corresponds to the upper left region of Figure 2.6 and 2.8 mostly affords the object 0; an agent tends to reach object 0. But scrutinizing these areas, we see that small gray dots are scattered in this region, and Figure 2.6 tells us that they are unreachable regions. This complex basin boundary is a characteristic of this agent. We interpret that different layouts afford different objects to an agent.

Also, it should be noted that the required time to get an object also forms complicated boundaries; there are mixtures of almost unreachable regions and relatively easy regions. On the other hand, striped regions lie in-between (-

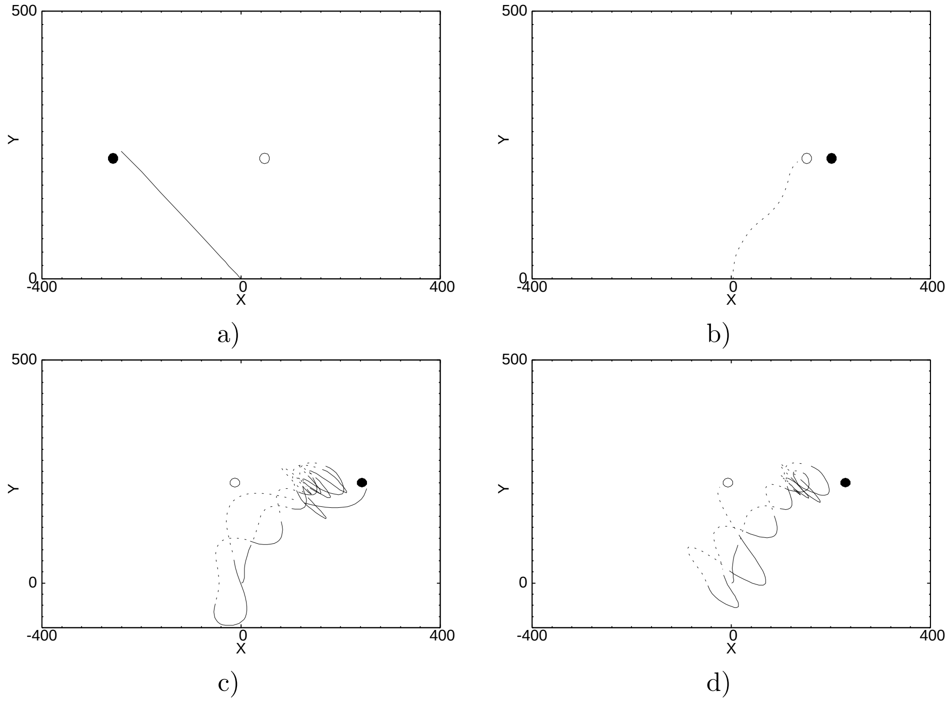


Figure 2.5: a) An example of the agent's straight line trail towards the object 0 for task 2. The coordinates of the objects are (-257.4,225) and (47.4, 225), and the time steps before reaching the object 0 is 22. b) An example of the agent's reaching the object 1, which takes 33 steps. The coordinates of the objects are (201.6,22) and (151.8,225). c) An example of spending many time steps before reaching an object 0 for the task 2. The coordinates of the objects are, (243,225) and (-12.6,225). A total time steps before reaching the object 0 is 376 and the associated entropy of the length 10 is computed as 0.372444 (see the definition of entropy in the text). d) An example of spending many time steps before reaching an object 1 for the task 2. The coordinates of the objects are (229.2,225) and (-6.6,225). A total time steps before reaching the object 1 is 326 and the entropy of the length 10 is 0.370865.

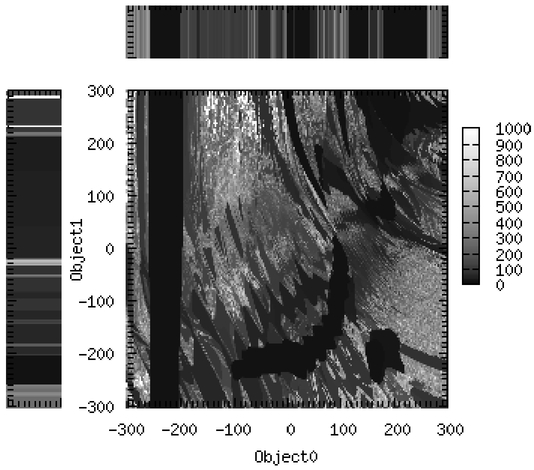
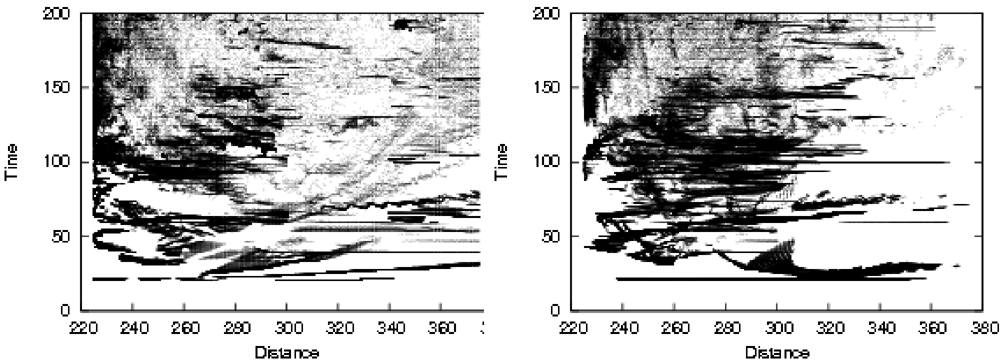


Figure 2.6: A grayscale plot of time steps required to reach an object in task 2. The darker area indicates less time steps is required to get an object. The x-coordinates of objects are taken as a horizontal and vertical axis, respectively. Two rectangular (a bar code) areas express the time steps of task 0 and 1.



a) b)

Figure 2.7: The required time distribution for reaching object 0 as a function of the distance from object 0 in task 0 and task 2. a) is in task 0. b) is in task 2.

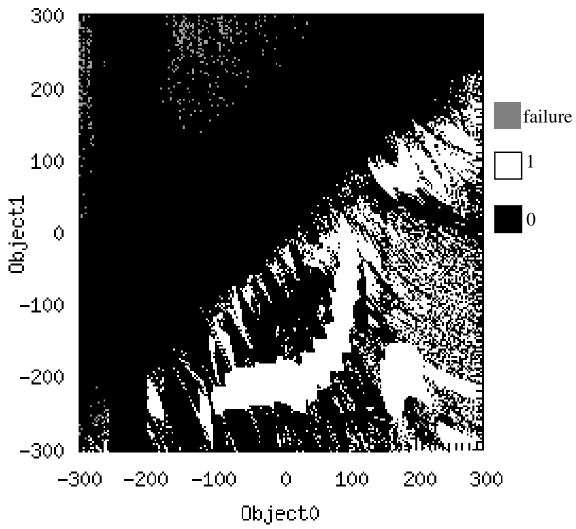


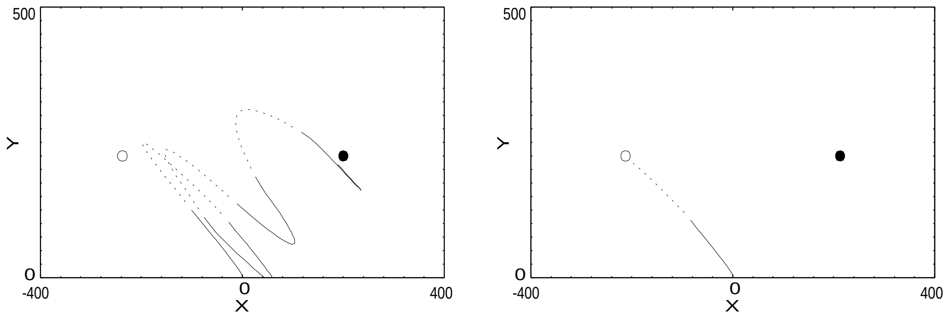
Figure 2.8: A basin structure for the selection of objects in task 2. A black region in this figure corresponds to object 0 and a white region does to object 1, and the gray region does to the un-selected region (none of the objects were selected).

250,-200) afford object 0 without taking more than 100 time steps to reach them. Also, the white crescent arch found in Figure 2.8 corresponds to the black crescent arch in Figure 2.6, which means that this region requires very few time steps to reach them, and no complex boundaries can be found in this region.

2.3.3 The difference of the layouts

Here we study more about the difference of similar layouts. It should be so that the similar layout must afford the same tendency. But this is not always the case (see Figure 2.5). For example, Figures 2.5 c) and d) have only a small difference, but their layouts afford different objects. In Figure 2.9, almost the same layouts afford different objects, and different action switching dynamics appear before reaching the objects. From these observations, we say that the action selection is sensitive to the fine layout structure. In particular, the epsilon neighborhood of the basin of selecting an object 0 can have points that belong to the basin of selecting object 1. This riddled nature of action selection is expected to characterize microslip phenomena.

The riddled nature of the complex boundary can be characterized by measuring the length of the boundary that separates the initial states of reaching objects 0 and 1. Figure 2.10 is a zoomed-in figure of the region from (0,100) to



a) b)

Figure 2.9: a) When the objects are at (200,225) and (-237,225), it takes 188 steps to get the object 1 and the associated entropy is 0.283701. b) For a slight different layout from a), where the objects are at (212.5,225) and (-212.5,225), the agent performs very differently. It only takes 28 time steps before getting the object and the entropy is 0.301057.

(100,0) in Figure 2.8. We see that we still have similar complicated regions. If this similar complex figure continuously appears by re-scaling the figure, we say that it has a fractal basin boundary. In particular, if the boundary dimension is the same as the space dimension, we call it a riddled basin structure. In order to measure the boundary length practically, we adopted the box counting method. Using this method, the boundary dimension of the action selection is approximated as around 1.761 (see Figure 2.11), which is not equal to the space dimension 2.0. We thus say that this does not form a riddled basin structure but has a fractal nature.

The fractal nature of the basin boundary confirms that the basin of action selection is indeed entangled, and the complexity of the action selection can be characterized by this, i.e., which to choose, action primitives of object 0 or object 1, is mostly undecidable. If the real microslip is characterized by this complexity of the boundary, we expect a fractal nature in the boundary.

2.3.4 Entropy of an action switching pattern

The other quantification of microslip is to use the entropy of the temporal sequence of action switching events. An agent achieves the goal by sequentially picking up the action primitives. Again, the action primitives are defined as a set of action fragments generated in task 0 and task 1. Defining two sets of action primitives as 0 and 1, we symbolize each temporal sequence as a binary string, such as 0101000101010000. Thus, the complexity of this binary string is measured in a straightforward way by the Shannon entropy(equation (2.6)).

$$E = -\frac{1}{N} \sum_{i=0}^{2^N-1} p_i \log p_i \tag{2.6}$$

In this experiment, we used $N = 10$ as the length of a bit string.

Here the $p(i)$ is computed as a occurrence probability of the pattern i. In this experiment, we adopt 10-bit length to calculate the entropy (the total number

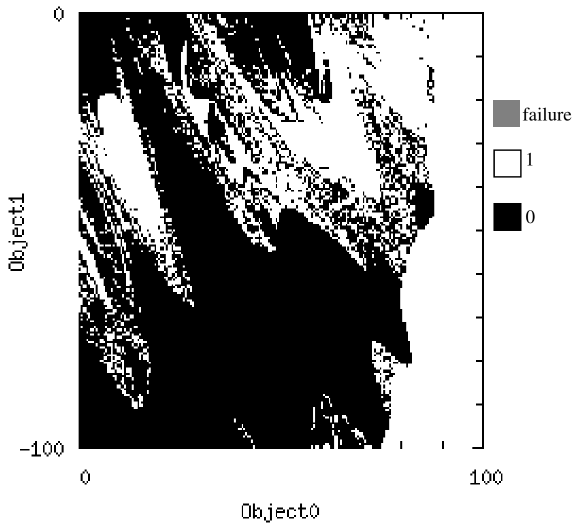


Figure 2.10: A scale up of Figure 2.8 of the range (0, -100)-(100, 0).

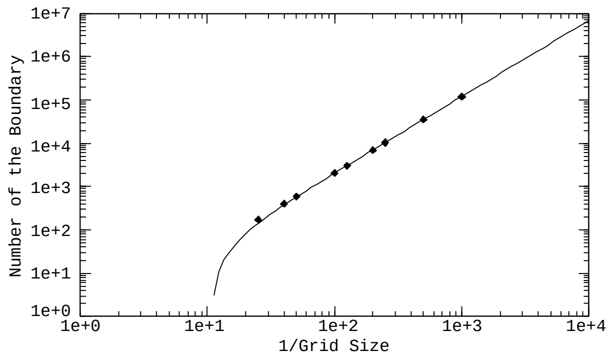


Figure 2.11: A computed Fractal dimension by the linear fitting $f(x) = 0.623x^{1.761} + C$

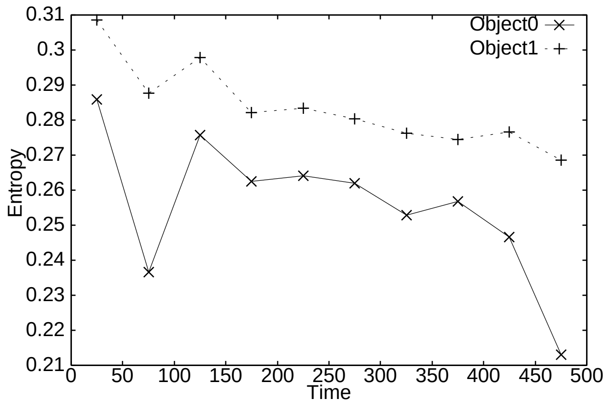


Figure 2.12: The entropy is computed separately for reaching objects 0 and 1 in task 2.

of the possible pattern is 1024) by bit shifting and averaging over the entire length.

Figure 2.12 is an example of task 2 where the agent spends 400-500 time steps to get an object. We computed the above entropy per every 50 time steps, which shows an abrupt decrease around 50-100 time steps, then gradually decreasing down to 0.21. This implies that the agent shows complex switching in the early stages, but the switching pattern becomes periodic in the later stages. In the end, the agent fails to get an object.

Figure 2.13 shows the entropy value (E) as a function of the number of switching times (Ns) for task 2. For example, the Ns of the string 0001101 are computed as 3. We say the string is complex when it has a large E with large Ns, and simple when it has a small E with small Ns. If the string has a small E with large Ns, we say that the string is periodic, but the opposite case is not seen frequently (i.e. large E with small Ns, in upper left side ($E > 0.3$ and $Ns < 0.05$) in each of Figure 2.13). Unfortunately, it happens that some short strings have a large E with small Ns. This is due to the artifact of our method, so we should be careful when dealing with the relatively shorter strings.

According to this definition, we classified the strings into 5 classes.

1) A complex string with a few time steps (i.e., $E > 0.3$ in Figures 2.13 a) and c). An actual orbit can be found in Figure 2.15 a)).

2) A periodic with a few time steps ($E < 0.3$ in Figures 2.13 a) and c) and the $Ns > 0.1$. An actual orbit can be found in Figure 2.15 b)).

3) A simple string with a few time steps. ($E < 0.3$ in Figures 2.13 a) and c) and the $Ns < 0.1$. An actual orbit can be found in Figures 2.5 a) and b)).

4) A complex string with long time steps. ($E > 0.3$ in Figures 2.13 b) and d). An actual orbit can be found in Figures 2.5 c) and d)).

5) A simple string with long time steps. ($E < 0.3$ in Figures 2.13 b) and d). An actual orbit can be found in Figure 2.15 c)).

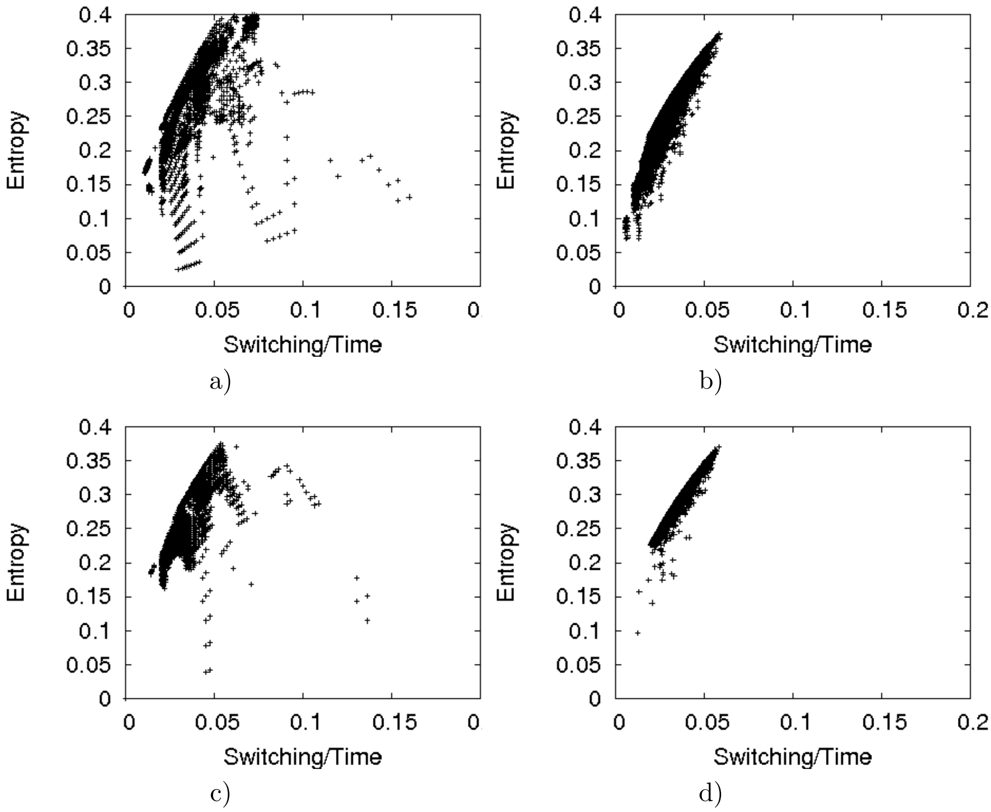


Figure 2.13: Entropy and Ns per each time step is calculated for task 2. a) An agent reaches the object 0 within 100 time steps. b) The agent spends 300-400 time steps to get the object 0. c) The agent reaches the object 1 within 100 time steps. d) The agent reaches the object 1 within 300-400 time steps.

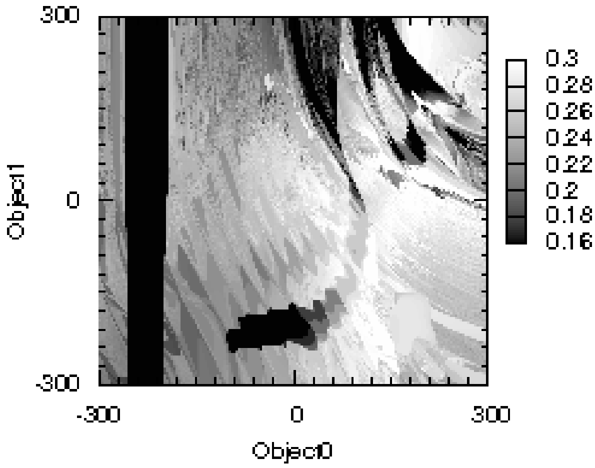


Figure 2.14: A diagram of entropy as a function of the layout pattern. The brighter area has the higher entropy values. To clarify the tendency, we make it black if the entropy is less than 0.16 and white if it is greater than 0.3.

Figure 2.14 computes the entropy as a function of the layout (the x-coordinates of objects 0 and 1). Comparing this diagram with Figures 2.6 and 2.8, we notice that whether the agent spends a longer or shorter time in reaching an object is not directly related to the entropy value. Also, we notice that the above classification 1)- 5) can be found in different portions of the diagram. For example, around the region (150,-100) - (250,0), the entropy is high, and the agent spends longer time steps in reaching an object. Also, the crescent arch in Figure 2.6 looks homogeneous, but the entropy of this region is gradually changing, as in Figure 2.14.

2.4 Discussion

Reed et al.(Reed et al., 1993) wrote in their unpublished paper that "actions are not made planful and purposive by the addition of a special conscious, attentive, mode of control, but emerge as planful purposive performances because of the intrinsic nature of the units of action." We agree with this view that the next action is produced each time an action is made, and the entire action sequence is generated in some ad-hoc way. By the term ad-hoc, we mean that the action is not something pre-determined by an action plan, but its execution is processed essentially in parallel ways, as (Gibson, 1966; Neisser, 1967; Reed, 1996) have been arguing.

No plan exists in advance, but as a result of action production, it reveals coherent intentionality on top of this ad-hoc dynamics. An intention to "make a coffee" is a good example of ad-hoc dynamics. The process of making an instant

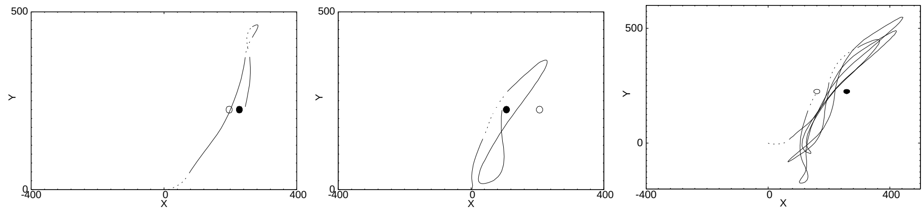


Figure 2.15: a) An orbit with a large entropy with a fewer numbers of switching. The coordinates of objects are at (226.8,225) and (196.2,225). It takes 69 steps and the entropy is computed as 0.40899. b) An orbit with a small entropy with a few number of switching. It takes 96 steps and the entropy is 0.13638. The objects are located at (106.2, 225) and (206.4,225). c) A small number of entropy with a fewer number of the switching behavior and longer time steps before reaching an object.The objects are at (258,225) and (160.2,225). It took 321 time steps and the entropy was 0.0967722.

coffee requires a chain of local actions (e.g., preparing some powder, bringing a cup, scooping up with a spoon, etc.), which is tangled with one's intention of drinking a cup of coffee.

In other words, an action has local and global temporal scales running at the same time. The present simulation shows that a global task is composed of the action primitives of objects 0 and 1; however, such a separation of primitives is not complete, so that produces the microslip phenomenon. Therefore, we have the following classification from the present dynamical systems approach.

A) Action stuttering within the same action primitive.
A large fluctuation exists within the same primitive. For example, Figure 2.15 c) displays the case where the agent goes back and forth without changing the action primitive type.

B) Action stuttering between two primitives.
A large fluctuation is observed in the switching behavior between two primitives.

A fluctuation within the same primitive (case B)) is generated by the instability of the fixed dynamics with inputs from the environment, whereas the other fluctuation (case A)) is caused by the switching dynamics, which can be referred to as an iterated functional system. Here we take case A) as an example of hesitation and B) as the other patterns. When the action hesitation (of case A)) occurs, it has a complex boundary pattern, and the other case B) has the simple boundary pattern. (But we have some exceptional cases here.) A subtle difference in layouts controls the microslip occurrence, which can be examined by a real psychological experiment. Two microslip classifications were analyzed with the entropy measure. The variety of switching patterns is characterized by the entropy; however, the relationship between the number of switchings and the entropy value has no apparent correlation.

- When the number of switchings per one-time step is proportional to the entropy, a case with the lower entropy corresponds to case A) and that of a case with higher entropy to case B).

- When it is not proportional to the entropy (more switching with the lower entropy case): cases A) and B) are mixed in this case. It is observed that there are more examples of case B) in the early stages and more examples of case A) in the later stages. In fact, a transition from case B) to A) is observed in Figure 2.12.

The relationship with other primitives is also learned by the network in addition to the primitives themselves; thus, the two primitives are not independent from each other. We should put stress on the last point. A combination of primitives and the formation of a hierarchy may be taken as characteristics of experimental microslips. But the action primitives are not independently prepared. They are self-organized by way of GA. In this simulation, we forced the network to learn three tasks. Task 2 is a combination of tasks 0 and 1. Therefore, the network must learn not only learn a primitive but also how to combine them. Namely, each module and the way to use those modules are expressed in the same network. The embedded patterns in one network is called "heterarchy" (i.e. a hierarchical network with connections in the same level) rather than a simple hierarchy (top-down like connections) one (e.g., (Jen, 2003; Nakajima et al., 2007)).

In our experiments, a basin structure of the final decision (0 or 1) shows that in some layouts, the epsilon neighborhood of the final decision 0 often contains the final decision 1 and vice versa, which organizes an almost riddled basin-like structure (Figure 2.8). We interpret that the corresponding layouts have an undecidable appearance, whereas the layout without such a complex boundary simply affords one object.

The complex basin structure has been well-known for non-linear systems (e.g., (Grebogi et al., 1986; Alexander et al., 1992)). (Nishimoto and Tani, 2004) studied a recurrent neural network (RNN) to let it learn a branching context represented by a finite state machine. They showed that an agent with an RNN can learn a future navigation "plan" in an initial phase space, where the fractal boundary is also observed. Their environment also contains some undecidable points, which is similar to our task 2 condition.

The above case A) suggests that a neural network without choice-controlling neurons (S_0, S_1) can perform the microslip phenomenon. Indeed, it would be interesting to study such microslips emerging in the non-modular type network where the internal neurons can spontaneously become the controlling neurons. Also we have used a discrete time evolution system here, but it would be interesting to test a continuous time recurrent network (CTRN). It should be noted that recurrent neurons in this network change their states discontinuously in time without taking intermediate values. Therefore, recurrent neurons certainly have a dynamic nature, but it is worth trying whether microslips can still emerge in neural nets without recurrent states. The necessary condition for causing microslips is left for future work; however, the combinatorial complexity arising from at least more than two intentionalities (case B)) provides a necessary condition in this study.

2.5 Summary

A microslip is an undistinguished feature of everyday life, but it gives us a deep insight into the principle of cognitive behavior behind action selection. In this paper, we have simulated microslips with a simple mobile agent with a discrete time recurrent neural network, and we have characterized it in terms of the complex basin structure and the entropy. An agent is always getting inputs from the environment. We investigated when an agent is sensitive to the spatial layout of objects rather to objects themselves. This notion of layout as an input may be useful when comparing with the cognitive experiments of affordance. The layout of two objects controls the agent's behavior. This is what we expect in real microslip experiments. Affordance of objects can not simply be attributed to the objects themselves but to the spatial layout. Actually, it is well-known that the change of layouts in a drugstore surprisingly changed consumers' behaviors (Cox, 1964). Some psychological experiments demonstrate that people extract different affordances from the same layouts (e.g., photos of landscapes) to find their way (Fukuma, 2003). In real situations, the layout of only two objects may be too simple to cause microslips; however, it is sufficient for understanding the essence of microslips.

Microslips are rooted in the singular phenomenon of dynamical systems (e.g., complex riddled basin-like phenomenon), which should be worth noting. Depending on the GA generation and the evolution run, microslips as this singular phenomenon appear or disappear (see the Appendix for GA history and different evolution runs). It is simple to ignore those agents that can produce singular behaviors. But we think this singularity is more important than having robots with stable behaviors, because this singular phenomenon is a characteristic of dynamical systems, which is not expected in finite state machines. Considering that the essential nature of cognition can be modeled with unstable dynamical systems (Ikegami, 2007; Tsuda, 2001), we inversely have to ask why we can have stable cognition in everyday life. In the case of microslips, there is a report that a patient who is suffering from brain damage and doing rehabilitation shows microslips more frequently (Sasaki, 2005; Reed et al., 1993). But by way of recovering from the damage, the patient can select an adequate action, and the frequency of microslips goes down (but not to zero frequency). A regulatory mechanism of selective attention or the function of awareness should be revealed both theoretically and experimentally. For the moment, we can say only that a microslip is a good example for studying this paradoxical conflict between the underlying singular dynamics and stable cognitive behaviors.

Chapter 3

Active Touch Feeling Evolved by Interactive Evolutionary Computation

3.1 Tactile sensation and experiments

People have been studying various aspects of the feelings of "touch" phenomena. Recently, tactile display technology has made great progress, and some new actuator designs have been proposed for creating realistic tactile sensation. For example, actuators with ultrasonic vibration (Watanabe and Fukui, 1995), piezoelectric elements (Pasquero and Hayward, 2003), electrical stimulation (Takahashi et al., 2002), and ICPF(Ionic Conducting Polymer gel Film) (Konyo et al., 2000) are the leading materials and ideas. Those actuators are designed to provide a mechanical vibration whose amplitude and frequency meet the human vibro-tactile threshold curve (see Figure 3.2). By activating those receptors, the corresponding neural circuits will be activated to deliver sensory feelings to a subject.

The question here is, whether the sensory feeling is determined only by the activation of receptors or by something else, e.g., actively controlled by a subject's "intentionality." Real tactile materials have complex patterns organized by bumps on the surface, the elasticity of yarn and fur, and so on. At the same time, human hand movement while touching also seems to have inherent dynamics, as Katz shown in Figure 1.1 and discussed in his book (Katz, 1925).

In this chapter, we try to investigate the nature of active perception on tactile sensation by using a human interactive genetic algorithm. We evolve a neural network that takes hand movement as an input and a vibration signal as an output. Studies on subjects' hand movements while touching objects have been conducted (Konyo et al., 2005; Sato et al., 2007). Influenced by previous studies, our work has the following characteristics:

1. The actuator consists of the ion conducting polymer gel film (ICPF) that vibrates smoothly due to the posed electric voltages.

2. The actuator output is the function of the hand movement.

3. The transformation from the hand movement (its velocity and acceleration) to the output is mediated by a recurrent neural network.
4. A structure of the neural network is selected by Interactive Evolutionary Computation (IEC).

In the following sections, we will explain how the experiment is organized and where we elaborated to make an accurate measurement.

3.2 Theories

3.2.1 Cutaneous mechanoreceptors and tactile sensation

Some mechanoreceptors of human skin have neural activity affected by vibration stimulus. The mechanoreceptors are classified into 4 types by a combination of adaptation speed(fast or slow) and receptive field(small or large) as shown below (Johansson and Vallo, 1983).

FA I Fast and small type. E.g., Meissner's corpuscle.

FA II Fast and large type. E.g., Paccinian corpuscle.

SA I Slow and small type. E.g., Merkel's disc.

SA II Slow and large type. E.g., Ruffini's ending.

The receptive field is defined as an area where a mechanoreceptor can respond to the given stimulus. FA means fast adapting, and SA means slow adapting response behavior. The I and II types imply a small and large receptive field, respectively. FA type receptors are activated while the stimulus is being put on and removed on the mechanical stimulus, and SA type receptors are activated while the receptors are being deformed. The 4 mechanoreceptors are placed at different depths measured from the skin surface as depicted in Figure 3.1.

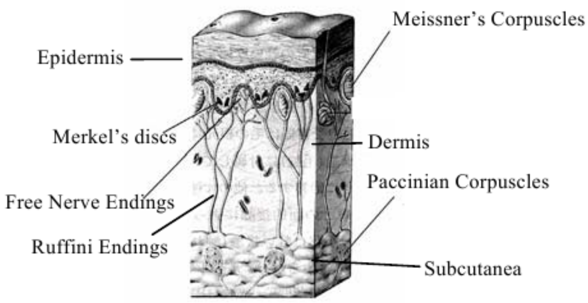


Figure 3.1: Structure of human skin tissue (from Maeno et al. (Mukaibo et al., 2005))

Figure 3.2 shows the neural discharge threshold of monkey mechanoreceptor units(FA I and II) and a human vibrotactile threshold curve. The threshold areas of the monkey FA I and II were obtained by Mountcastle et al. (Mountcastle et al., 1972), and the threshold curve of human FA I and II were obtained by Miyaoka et al (Higashiyama et al., 2000). The human curve is composed by the human reports. The monkey FA I optimally responds in the range of 10-50 Hz, and FA II does in that of 100-300Hz, as shown in the figure. The monkey FA I and II areas have some coincidence with the human response curve. Therefore, Miyaoka et al. claimed that the response performance of human receptors is similar to those of a monkey (Higashiyama et al., 2000).

On the other hand, Bolanowski et al. measured the vibrotactile threshold curves of SA I, and claimed that SA I responds stronger than FA I under 2 Hz of stimuli (Bolanowski Jr. et al., 1988). Moreover, the slope of the curve under 20 Hz in Figure 3.2 is approximately -1, but the slope at 50-300 Hz is approximately -2. From this observation, Miyaoka et al. claimed that the vibrotactile threshold curve in Figure 3.2 is composed of at least 2 types of a system (Miyaoka et al., 1985).

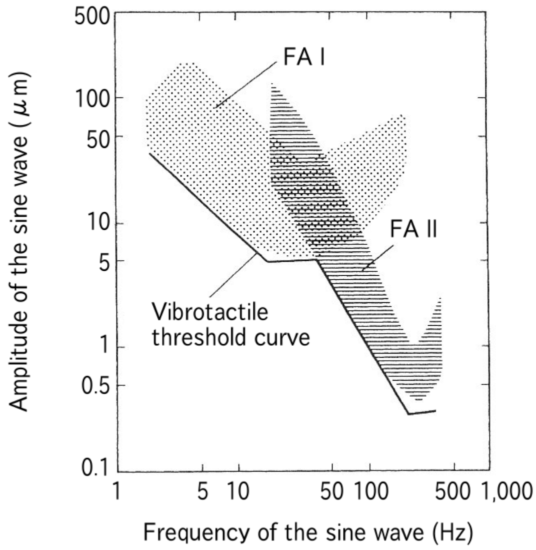


Figure 3.2: Neural discharge threshold curves of monkey mechanoreceptor units and the human vibrotactile threshold curve (revised from Miyaoka (Higashiyama et al., 2000) which is revised from Mountcastle et al. (Mountcastle et al., 1972))

In terms of the classification of the functions of tactile sensation, Miyaoka et al. discussed three kinds of perception: subtle stimulus, fine surface texture, and rough surface texture pattern (Higashiyama et al., 2000). First, the subtle stimulus is caused by, e.g., movement of a tiny insect on the skin. People can detect an extremely subtle stimulus on a large area of their skin. Miyaoka et al. also claimed that the detection of such a subtle stimulus is mainly attributed to the property of FA II, because FA II has a large receptive field with fast adaptation.

Second, the fine surface texture is, for example, an abrasive paper that has several 10 μm particles. People can better perceive it by using their fingertips. Miyaoka et al. claimed that it is mainly caused by FA I, because the density of FA I on skin is higher in the fingertips than in other areas (Miyaoka et al., 1999).

Last, the rough surface texture is, for example, a visible rugged floor or cloth. Johnson et al. investigated the relationship between a subjective roughness feeling and the spatial distribution of mechanoreceptors on the skin, and the researchers claimed that SA I receptors are responsible for that sensation (Blake et al., 1997).

When mechanoreceptors are stimulated by vibration, they are adapting to the stimulus. Hahn investigated the vibro-tactile adaptation of fingers and its recovery by putting a 60 Hz vibration (Hahn, 1966), and discovered that the time duration required for recovery is half its adaptation.

3.2.2 ICPF tactile display

We use a tactile display developed by Konyo et al. (Konyo et al., 2000) as shown in Figure 3.3. The display has ICPFs (Ionic Conducting Polymer gel Films) with gold plating as an actuator to stimulate fingers. Figure 3.4 shows the architecture of the ICPF actuator. When an electric field is applied, ions are transported to one side of the film, and it is bent. The film oscillates swiftly due to the oscillatory electric field. Comparing this actuator with the other actuators, we notice that the advantage of this ICPF actuator is the high-speed response and the low driving voltage. The response speed is more than 100 Hz, and the driving voltage is less than 3 V. However, since the inside of the ICPF is filled with water in order to drive ion currents, this tactile display is sometimes soaked in water during the experiments.

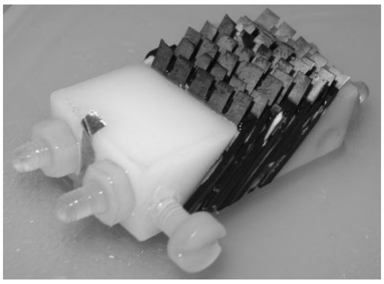


Figure 3.3: A tactile display using ICPFs, which are rectangular slices. Subjects wear it on the finger cushion of their index finger.

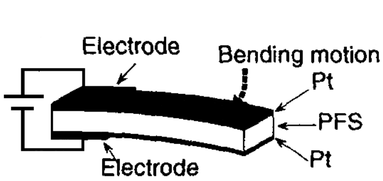


Figure 3.4: ICPF actuator (from Konyo et al. (Konyo et al., 2000))

3.2.3 Interactive Evolutionary Computation (IEC)

To emulate subjective tactile sensation like a real feeling, we use Interactive Evolutionary Computation (IEC). IEC is a technique in which the system is optimized by using a human choice as a fitness function, and originated with Dawkins (Dawkins, 1986) (and by some other people). Recently, IEC has been used for many other purposes (Takagi, 2001). Here, we use neural networks as an "learning machine" for obtaining tactile sensation. As an application of IEC to evolve neural networks, Lund et al. used it to move robots (Lund et al., 1998).

3.2.4 Setting up an evolutionary target

First, we tested our method to emulate simple touch feeling, and then used Japanese onomatopoeias to use human subjects to develop the sensations. We started by using the touch feeling of hemp as a goal of the experiment (Ogai and Ikegami, 2007). From the experiment, we noticed that each subject evolved very different tactile sensations with different neural networks. Each subject may pick up different characteristics of the feeling, such as the unevenness and smoothness of hemp to emulate the sensation.

As is discussed in the work of Ikegami and Zlatev (Ikegami and Zlatev, 2008), onomatopoeia can be a precursor to language and is a highly embodied system. Researchers have also discussed that onomatopoeia represents emotion more directly than do other words (Osaka, 1999). Since onomatopoeia is grounded in embodiment, we use onomatopoeia in this experiment to see the association between the onomatopoeia and hand movements. We asked subjects to generate the sensation indexed by onomatopoeia. Practically, we performed the experiment with 2 different types of onomatopoeia, *uneune* and *zarazara*, which we will explain in 3.4.2.

3.3 Development of the system

3.3.1 Outline of the system

We show an outline of the system for our experiment in Figure 3.5. Figure 3.6 is a picture in which a subject wears a sensor and the ICPF display on his hand. The sensor is put on the back of his hand, and the display is under the subject's index finger. A chart of the experiment is as follows. Repeating the 6 steps during every round of IEC evolved a given neural network.

1. A subject wears the ICPF display and moves his hands.

2-a. A 3D spatial position of the subject's hand position per each 16.67 msec, which is detected by a 3D magnetic position sensor, is sent to a PC.

2-b. The PC calculates the hand's velocity and acceleration from the data. A recurrent neural network (RNN) running in the PC uses the calculated velocity and the acceleration as the inputs, and the outputs are the calculated voltages.

2-c. The calculated voltages go through an amplifier to the ICPF display. The display adds vibrotactile stimulus to the subject's finger.

3. The subject judges the produced sensation by repeating steps 2-a to 2-c.

4. Based on step 3, a better RNN will be selected, and one will be replaced with a better one on some modifications in the weights.

The clock frequency of the 3D magnetic position sensor is set at 60 Hz. Okamoto et al. reported that a time-delayed response shorter than 40-60 ms will not be detected (Okamoto et al., 2008). Therefore, we designed the experiment so that step 2 occurs every 16.67 ms, that is, 60 Hz. The 3D magnetic position sensor will be described in more detail in Section 3.3.3. RNNs will be described in detail in Section 3.3.4, and the software and devices around the PC will be described in Section 3.3.7. The whole IEC process will be described in 3.3.5.

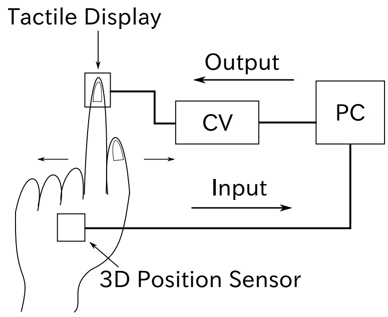


Figure 3.5: An outline of the system

Figure 3.6: The sensor and the ICPF tactile display on a hand

3.3.2 Output voltage

FA I, FA II, and SA I receive and transmit the sensation of the texture stimulus as described in Section 3.2.1. FA I and II respond to the low and high frequencies, respectively. Therefore, we use vibrations of 30 Hz and 180 Hz to stimulate FA I and FA II. SA I responds well under 2 Hz of vibration. The current system, however, works on 60 Hz, which is much faster than 2 Hz. Therefore, we will not be able to effectively stimulate SA I explicitly in this experiment. However, we expect the evolved RNNs may emulate the lower frequency so that it can also stimulate SA I. Thus, we adopted Equation (3.1), which is a combination of sinusoidal functions with 2 different frequencies (f_1(30 Hz) and f_2(180 Hz)). $V(t)$ is the output voltage for the ICPF display.

$$V(t) = \begin{cases} A_1(n)sin(2\pi f_1 t) + A_2(n)sin(2\pi f_2 t) & (Ve(n) \geq K_{th}) \\ 0 & (Ve(n) < K_{th}) \end{cases} \tag{3.1}$$

We use the following notations: t is time, Δt is an update interval, $A_1(n)$ and $A_2(n)$ are the amplitudes of sinusoidal waves, n is the integer part of $[\frac{t}{\Delta t}]$, $Ve(n)$ is the hand's velocity, and K_{th} is the threshold to set $V(t)$ to zero when the hand's velocity is too slow. $Ve(n)$ will be described in detail in Section 3.3.3. $A_1(n), A_2(n)$ are calculated from outputs of the RNN, and will be described in detail in Section 3.3.4. Here, K_{th} is 0.1, a unit of t and Δt is second, and Δt is 1/60 seconds because of the clock frequency of the sensor. We set the maximum limit of $V(t)$ to 3 V, because of the performance of the ICPFs. Therefore, $A_1(n)$ and $A_2(n)$ are adjusted. They will be described in detail in Section 3.3.4. In all equations after this section, t, n, and Δt will be used with the same definitions as in this section, because the systems represented by the equations are synchronized with each other.

3.3.3 Inputs from the 3D magnetic position sensor

The 3D magnetic position sensor is "Polhemus Patriot" as shown in Figure 3.7. Its clock frequency is 60 Hz, and it detects 6 degrees of freedom. We use $X(n)$ and $Y(n)$ parameters only in the horizontal direction. A unit of $X(n)$ and $Y(n)$ corresponds to about 2.5 cm in real space. The hand's velocity Ve and acceleration Ac are given by

$$\begin{aligned} Ve_X(n) &= X(n) - X(n-1) \\ Ve_Y(n) &= Y(n) - Y(n-1) \\ Ve(n) &= \begin{cases} K_{Ve}\sqrt{Ve_X(n)^2 + Ve_Y(n)^2} & (Ve(n) \leq 1) \\ 1 & (else) \end{cases} \\ Ac(n) &= \begin{cases} K_{Ac}\sqrt{(Ve_X(n) - Ve_X(n-1))^2 + (Ve_Y(n) - Ve_Y(n-1))^2} \\ \qquad (Ac(n) \leq 1) \\ 1 \qquad (else) \end{cases} \end{aligned} \tag{3.2}$$

where, K_{Ve} and K_{Ac} are the coefficients to adjust $Ve(n)$ and $Ac(n)$, respectively. $Ve_X(n)$ and $Ve_Y(n)$ are the intermediate variables to calculate $Ve(n)$ and $Ac(n)$. Here, K_{Ve} is 10, and K_{Ac} is 30, because we set the maximum velocity as approximately 15 cm/s, and the maximum acceleration as approximately 45 cm/s^2. We decided the values of the maximum velocity and acceleration based on the speed of the author's hand.

3.3.4 Recurrent Neural Network (RNN)

We show the RNN in Figure 3.8. It has 3 layers. The first lowest layer gets Ve, Ac, and recurrent states as parameters. Initially, the recurrent parameters are

Figure 3.7: Polhemus Patriot (from Polhemus's website (Polhemus,))

set to 0.5. The states of the second and third layers' units $u_{i,j,k}(n)$ are updated according to Equation 3.3. Here, i is the number of a layer, and j is the number of a unit on the layer i. To see the dynamics of neural dynamics, we iterate the neural network M times for Δt, where k indexes the number between 0 and M.

$$u_{i,j,k}(n) = \begin{cases} g(\sum_{l=0}^{N_{i-1}} w_{i-1,j,l} u_{i-1,l,k}(n)) & (u_{i,j,k} \neq a_{1,k} \cap u_{i,j,k} \neq a_{2,k}) \\ \sum_{l=0}^{N_{i-1}} w_{i-1,j,l} u_{i-1,l,k}(n) & (u_{i,j,k} = a_{1,k} \cup u_{i,j,k} = a_{2,k}) \end{cases} \tag{3.3}$$

where, $w_{i-1,j,l}$ is a weight between $u_{i-1,l,k}(n)$ and $u_{i,j,k}(n)$, N_{i-1} is the number of units on the layer $i-1$, and $a_{1,k}(n)$ and $a_{2,k}(n)$ are outputs of the RNN. Here, $w_{i-1,j,l}$ is given by $-1 \leq w_{i-1,j,l} \leq 1$.

g(x) is a sigmoid function and given by

$$g(x) = \frac{1}{1+e^{-2x}}. \tag{3.4}$$

This function is not applied for $a_{1,k}$ and $a_{2,k}$.

$A_1(n)$ and $A_2(n)$ in Equation 3.1 are given by

$$\begin{aligned} A_1^{\prime}(n) &= \frac{1}{M}\sum_{k=0}^{M-1} a_{1,k}(n) \\ A_2^{\prime}(n) &= \frac{1}{K_{A2}M}\sum_{k=0}^{M-1} a_{2,k}(n) \\ A_1(n) &= \begin{cases} A_1^{\prime}(n) & (A_1^{\prime}(n)+A_2^{\prime}(n) \leq V_{max}) \\ \frac{V_{max}A_1^{\prime}(n)}{A_1^{\prime}(n)+A_2^{\prime}(n)} & (A_1^{\prime}(n)+A_2^{\prime}(n) > V_{max}) \end{cases} \\ A_2(n) &= \begin{cases} A_2^{\prime}(n) & (A_1^{\prime}(n)+A_2^{\prime}(n) \leq V_{max}) \\ \frac{V_{max}A_2^{\prime}(n)}{A_1^{\prime}(n)+A_2^{\prime}(n)} & (A_1^{\prime}(n)+A_2^{\prime}(n) > V_{max}) \end{cases} \end{aligned} \tag{3.5}$$

where, M is the execution time per Δt seconds, V_{max} is the maximum limit of $V(t)$, and K_{A2} is the coefficient to suppress the amplitude of the sinusoidal wave of f_2. Since FA II is much more sensitive than FA I, we have to balance the amplitude by K_{A2}. $A_1'(n)$ and $A_2'(n)$ are the intermediate variables. $A_1(n)$ and $A_2(n)$ are normalized when their summations are over V_{max}, respectively. Here, M is set at 100, V_{max} is set at 3, and K_{A2} is set at 2.

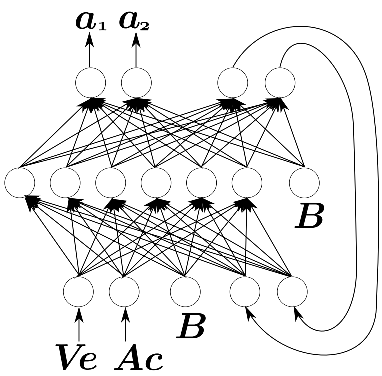


Figure 3.8: RNN used in the experiment. Ve: Velocity of a hand, Ac: Acceleration of a hand, a_1, a_2: Outputs to calculate A_1, A_2 in Equation 3.5, respectively, B: Bias neuron

3.3.5 Interactive Evolutionary Computation (IEC)

Let us discuss the IEC in more detail. The IEC here consists of 2 steps.

1) A subject examines two candidates to judge which one mimics the feeling of the given onomatopoeias by moving his hands better.

2) A better network replaces another network with some modifications (by a normal distribution of 0 mean and 0.1 deviation) on its weight values.

Figure 3.9 shows the chart. The above 1) and 2) will be repeated until the subject continuously selects the same network 10 times. We asked subjects to do IEC from the same neural weights in every trial; therefore, we expect that the sensations produced by the network will gradually converge as illustrated in Figure 3.10, where the space is an image of the network weights.

3.3.6 Tactile display

Subjects wear the tactile display under their left index fingers as shown in Figure 3.6. They also wear a guide of electric cords of the display and the sensor on the left elbow. The electric cords of the display go through a palm and are guided to the amplifier. Thus, the cords are adjusted not to disturb the subjects and the system. The display is held by a clear plastic cover, and pressed into the fingers by the elasticity of a pink sponge in Figure 3.6. The cover is attached with an adjustable hook-and-loop fastener. However, the cover and sponge are not sometimes enough to hold the fingers. Therefore, the subjects support the display with their thumb, if needed. When the ICPFs are used for a long time, the ICPFs' power gradually weakens due to the loss of water inside. Therefore,

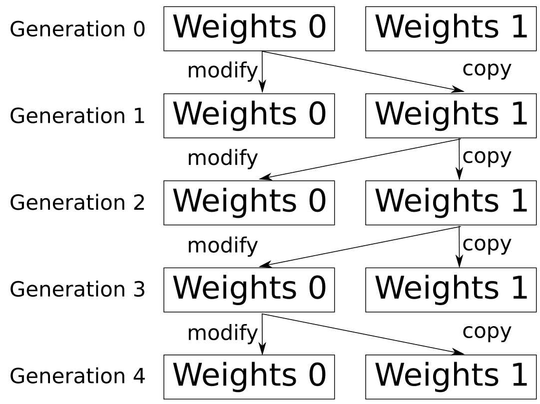


Figure 3.9: Evolution of weights

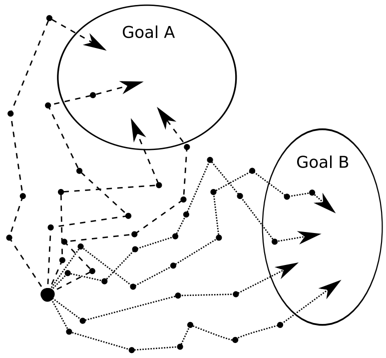


Figure 3.10: Image of evolutions

during the experiment to evolve networks, the subjects can soak the display in water whenever they feel that the ICPFs' power is weak, and during the experiments that we conduct by using the evolved networks, we soak the display in water long enough to refresh the ICPF before each trial.

To check whether the ICPF actuators correctly respond to the input voltage, we measured the vibrotactile thresholds by using 3 subjects, including the author, with the display by the stair-case method that determines a threshold by changing the amplitude of the stimulus gradually and exploring the relevant stimuli around a candidate of the threshold to double-check whether the threshold is right or wrong (Cornsweet, 1962). We applied stimuli of 30, 60, 90, 150, and 240 Hz 10 times each in random order (Figure 3.11). If the number of subjects and the frequencies are few, the down trend is similar to Figure 3.2. From the result and the author's experience, we adopted 30 Hz and 180 Hz as the main frequencies for the display.

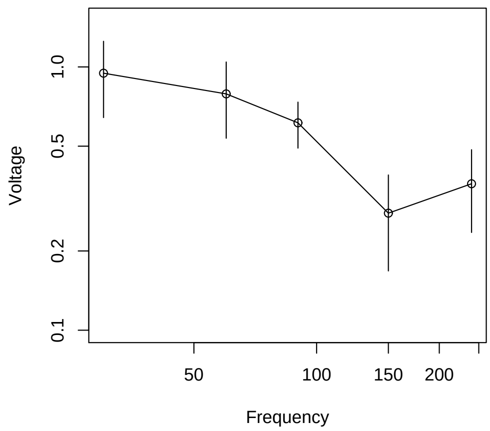


Figure 3.11: Means and error bars of vibrotactile thresholds

3.3.7 Software and devices around the PC

We developed software for the PC (NT9500pro of Epson) by using LabVIEW 8.6 of National Instruments on Windows XP SP3. Some other details are as follows: to output data from the PC, we used DAQCard-6062E of National Instruments in PCMCIA of the PC, and the DAQCard-6062E is connected to the amplifier via a BNC adapter (BNC-2110 of National Instruments). The amplifier is a "Bipolar Power Supply/AMP BWS 40-7.5" made by Takasago, LTD. Japan, and used to produce a constant voltage.

LabVIEW 8.6 and Windows are not for precise real-time experiments, but we adjusted them to keep the time delay under a few milliseconds. For example, the main computational loop is executed every 16.67 ms (i.e., 60 Hz). Figure 3.12 shows an example of the interval time of the main loop. We notice the interval time delays are kept under a few milliseconds. Additionally, the sampling rate of the output voltage is set at 1/10000 seconds.

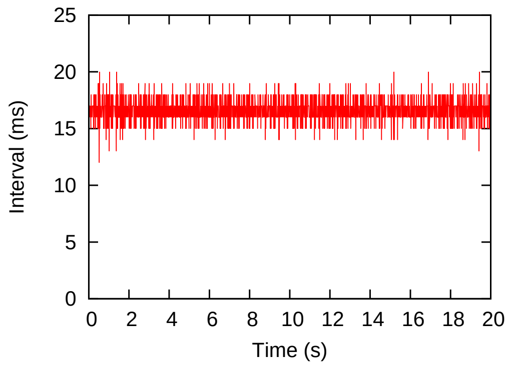


Figure 3.12: Interval of the main loop in the software

3.4 Experiments

3.4.1 Outline of the experiments

We conducted 6 types of experiments as shown below for each subject.

Experiment 1: Evolution of sensations

Experiment 2: Distinction tests between sensations evolved as *uneune* and *zarazara*

Experiment 3: Distinction tests between sensations that emerge from a subject's hand movement and the dummy movement generated by sine waves

Experiment 4: Distinction tests between sensations evolved by a subject who takes this test and other subjects

Experiment 5: Measurements of noise thresholds

Experiment 6: Writing sentences using *uneune* and *zarazara*

In Experiment 1 written above, a subject evolves the sensations generated by the weights of the RNN in order to represent 2 onomatopoeias (*uneune*, *zarazara*). The 2 onomatopoeias will be described in detail in Section 3.4.2. After the evolving the RNN, in order to collect data about a subject's hand movement and the dynamics of the neural networks, a subject uses the system with his hand movements to feel the onomatopoeias for 20 seconds. Next, we conduct distinction tests in which a subject distinguishes between the evolved RNNs and others in Experiments 2, 3, and 4. The distinction tests will be described in detail in Section 3.4.3. In Experiment 5, a subject measures the noise thresholds of sensations he evolved. The noise threshold measurements will be described in detail in Section 3.4.4. Additionally, just after Experiments 1, 2, 3, and 4, a subject answers questionnaires. The questionnaires will be described in detail in Section 3.4.5. Last, a subject writes sentences using *uneune* and *zarazara* in Experiment 6.

In Experiment 1, all subjects start from common initial weights shown in Table 3.1, and a common random seed. The initial weights are selected from some weights generated by random numbers by the author.

Figure 3.13 is the environment of the experiments. The system is controlled by a PC mouse through the interface of the software as shown in Figures 3.14, 3.15, and 3.16. The subject controls the software by himself. The subject uses the mouse with his right hand, and wears the display on his left index finger.

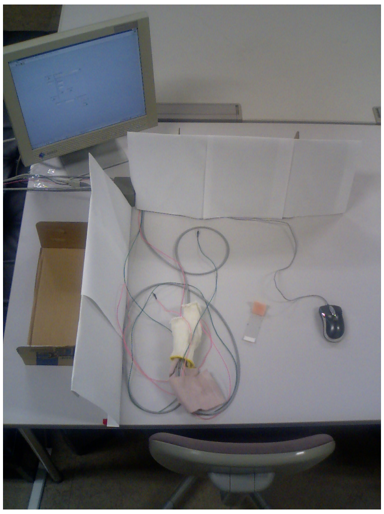


Figure 3.13: Environment of the experiment

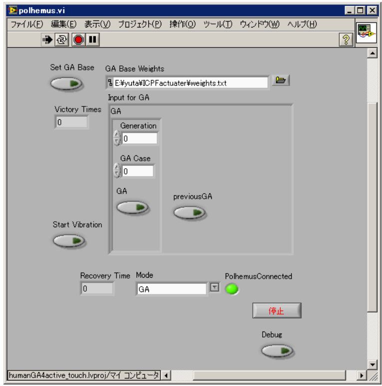


Figure 3.14: The window of the software used in Experiment 1

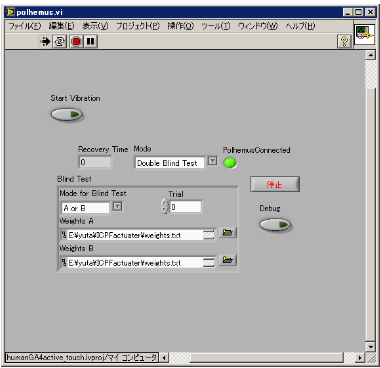


Figure 3.15: The window of the software used in Experiments 2, 3, and 4

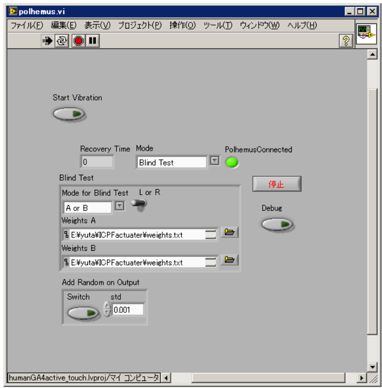


Figure 3.16: The window of the software used in Experiment 5

The subjects in this thesis are 2 men and 2 women, and their ages are 20-30s. All they are native Japanese speakers. We call the 4 subjects Subject A, B, C, and D. Subject A is the author. When Subjects A and B first evolved weights, noises added input. Therefore, the 2 subjects evolved the weights again from the weights evolved in last time.

Table 3.1: Initial weights

2nd layer \ 1st layer	Recurrent 0	Recurrent 1	Ve	Ac	Bias
Hidden 0	0.210221	0.245573	0.490896	-0.193924	-0.775185
Hidden 1	0.141132	-0.886595	-0.896890	-0.104405	0.700603
Hidden 2	-0.788986	0.837709	-0.017819	-0.272515	-0.846736
Hidden 3	-0.108717	0.672795	-0.686187	0.087836	0.771067
Hidden 4	0.472193	0.089273	0.853041	0.133574	0.908699
Hidden 5	0.295213	-0.134456	0.405703	-0.226785	0.700610

3rd layer \ 2nd layer	Hidden 0	Hidden 1	Hidden 2	Hidden 3	Hidden 4	Hidden 5	Bias
Recurrent 0	-0.171757	-0.226988	-0.906732	-0.165037	-0.575033	0.170103	0.127789
Recurrent 1	0.693498	0.874693	-0.879634	0.424730	-0.683270	0.290369	-0.378955
a_1	-0.169234	-0.222078	-0.213211	-0.329551	-0.259546	-0.727746	-0.834364
a_2	-0.497318	0.549449	0.409335	-0.037400	0.193336	-0.745483	-0.322009

3.4.2 Onomatopoeias as evolutionary goals

We use *uneune* and *zarazara* as target onomatopoeias. The 2 onomatopoeias are Japanese words. *Uneune* means "the touch of winding things," and *zarazara* means "the touch of a coarse surface." Namely, the onomatopoeias represent different types of tactile sensation. We believe that real tactile materials such as fur, corduroy, and others have internal dynamics due to the state of the material such as bumps on the surface, hysteresis as a condition of the hair, elasticity, and others. Moreover, we also consider that onomatopoeias such as *uneune* have similar dynamics to the condition of hair and the elasticity of the materials. Therefore, we make RNNs obtain the ability to represent the internal dynamics.

3.4.3 Distinction test

We conduct distinction tests in Experiments 2, 3, and 4. A subject distinguishes between 2 sensations in the experiments. Each test in the experiments has 20 trials composed of 2 different kinds of sensations in random order. The system can move for 20 seconds from when a subject starts the system by himself, and a subject can stop the system at anytime. A subject can try a trail that he already tried as many times as he wants.

The order of the trials is randomly chosen by software. Before the tests start, a subject can touch the 2 sensations that will be used in the test for as long as he needs. A subject controls the software as shown in Figure 3.15 by himself.

$P(k)$, which is the probability when the number of mistakes does not exceed k, is given by Equation 3.6:

$$P(k) = \frac{1}{2^{20}} \sum_{i=0}^{k} ({}_{20}C_i). \tag{3.6}$$

Then, $P(5) = 0.0207$, $P(6) = 0.0577$, and $P(7) = 0.1316$. However, when a subject makes too many mistakes in the tests, we consider that he distinguishes between them. In short, we adopt a two-tailed test. $P\acute{(k)}$, which the probability when the number of mistakes is not lower than k, is given by Equation 3.7:

$$P\acute{(k)} = P(20 - k). \tag{3.7}$$

Here, we choose 5% as the significance level. Therefore, $k = 5$ and $k = 15$ are within the significance level.

In Experiment 2, a subject distinguishes between his evolved sensations of *uneune* and *zarazara* in the distinction test. If a subject succeeds in evolving 2 different type of sensations, he can distinguish between them in the distinction test.

In Experiment 3, a subject distinguishes between sensations that emerge from his hand movement and the dummy movement generated by sine waves to investigate the contribution of the hand movement to the sensations. The dummy movement is defined based on the back-and-forth motion of 3 Hz. Here, the velocity $Ve(n)$ and the acceleration $Ac(n)$ of the dummy movement are described by

$$\begin{aligned} Ve(n) &= K_{amp}(sin(6n\Delta t\pi) + 1) \\ Ac(n) &= K_{amp}\frac{K_{Ac}}{K_{Ve}}(Ve(n+1) - Ve(n) + 2K_{amp}sin(3\Delta t\pi)) \end{aligned} \tag{3.8}$$

where K_{amp} is the amplitude. $2K_{amp}sin(3\Delta t\pi)$ in Equation 3.8 is added to set $Ac(n)$ over 0. The reason is given by

$$\begin{aligned} Ve(n+1) - Ve(n) &= K_{amp}(sin(6(n+1)\Delta t\pi) - sin(6n\Delta\pi)) \\ &= 2K_{amp}cos(\frac{6(2n+1)\Delta t\pi}{2})sin(\frac{6\Delta t\pi}{2}) \\ &\geq -2K_{amp}sin(3\Delta t\pi). \end{aligned} \tag{3.9}$$

The amplitude K_{amp} is 0.3 V. K_{amp} and the 3 Hz of the back-and-forth motion were decided based on the author's hand movements. Here also, if the velocity of a subject's hand movement is less than the threshold K_{th}, $V(t)$ is set to 0.

In Experiment 4, a subject distinguishes between his evolved sensation and other subjects' evolved sensations to investigate difference of their sensations. Each subject tried the sensations of other 3 subjects. Therefore, the combinations are $4 \times 3 = 12$ for each onomatopoeia.

3.4.4 Noise threshold

In Experiment 5, the subject measures the noise thresholds of his evolved sensations. In this experiment, the normal distribution of random numbers is added as noise to the outputs $V(t)$ of sensations. The noise threshold is a value that when the standard deviation of the noise is less, a subject feels the same sensation as when no noise is added. However, when the noise is more than the noise threshold, a subject feels the other sensation. The subject finds the noise threshold by Method of Adjustment in which a subject finds a value by controlling it by himself (e.g., (Gescheider, 1997)). Here, a subject controls the standard deviation of the noise. A subject uses the software as shown in Figure 3.16 to find the noise thresholds.

3.4.5 Questionnaire

We use questionnaires to ask a subject about his impressions after Experiments 1, 2, 3, and 4. We use Likert scaling (Likert, 1932) in the questionnaires. In Likert scaling, a subject chooses 1 from the 5 level items as shown below.

1. strongly agree
2. agree
3. undecided
4. disagree
5. strongly disagree

We use 3 types of questions that are written in Japanese. English translations of the questions are shown below.

(1) The evolved sensation corresponded to the goal onomatopoeia.

(2) You felt a difference between the 2 sensations.

(3) When you thought that you were touching the sensation moved by the dummy inputs, you felt that you were moved.

(1) is used after Experiment 1, (2) is used after Experiments 2, 3, and 4, and (3) is used after Experiment 3.

3.5 Results

3.5.1 Evolutions of sensations

We show the scored final products in Experiment 1 in Tables 3.2 and 3.3. The "Generation" column shows the number of generations spent for the evolution. The "Times of Change" column shows how many times a subject switches to a new weights until the last generation. The "Evaluation" column shows an answer to the question (1) as described in Section 3.4.5.

The values of Subjects A and B in the "Generation" column are bigger than the others. The reason seems to be that the 2 subjects evolved weights again from the weights that evolved with noises as described in Section 3.4. The "Evaluation" column of *zarazara* of Subjects B and C has 3 that are "undecided," and the other 2 are "agree." These results indicate the subjects generally succeed in evolving the weights.

Table 3.2: Evolution of *uneune*

Subject	Generation	Times of Change	Evaluation
A	60	10	2
B	111	30	2
C	51	12	2
D	30	7	2

Table 3.3: Evolution of *zarazara*

Subject	Generation	Times of Change	Evaluation
A	73	18	2
B	109	28	3
C	39	7	3
D	41	11	2

Figure 3.17 shows the evolution of weights measured by the Euclidean distance from the initial weights while the subjects evolved the weights. The horizontal axis of the figure is the distance of the lower layer weights, and the vertical axis of the figure is the distance of the higher layer weights. The distances gradually proceed diagonally right up, because normal distribution random numbers

change the weights. However, the movements of the distances seem to depend not only on random numbers but also on the judgments of the subjects, because, for example, *zarazara* of Subject A returns on the way.

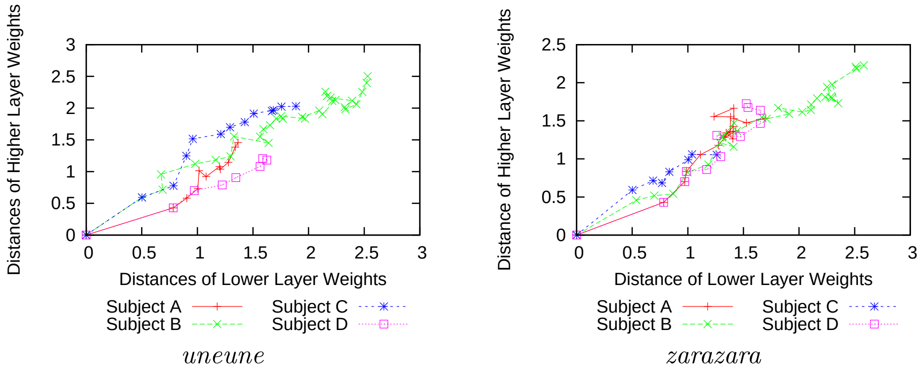


Figure 3.17: Euclidean distances of weights while they are evolved

Figure 3.18 shows movements of the averages of the RNNs' outputs $A_1(n)$ and $A_2(n)$ while the subjects evolve the weights of the RNNs. The black circles in the figure are the starting points, and the circles in each color are the end points. The points do not go straight, but wander. The movements of the points are not similar to the image of the evolutions in Figure 3.10.

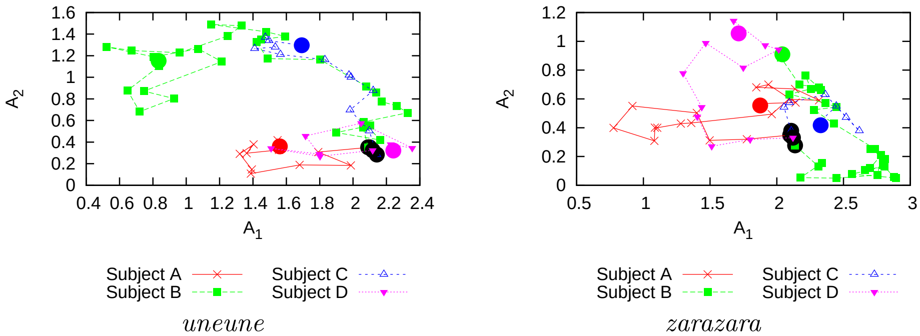


Figure 3.18: Averages of the RNNs' outputs $A_1(n)$ and $A_2(n)$ while evolving the weights of the RNNs

3.5.2 Patterns of RNNs' outputs and voltage outputs

After the evolutions, the subjects feel the sensation of their evolved weights again in order to collect data for 20 seconds as we have described in Section 3.4. Figure 3.19 shows the averages of the RNNs outputs $A_1(n)$ and $A_2(n)$ that are produced then. The points of the figure are similar to the end points of Figure 3.18 respectively. *Uneune* and *zarazara* are almost divided into right and left except for the *uneune* of Subject D. Moreover, *zarazara* of Subject C and *uneune* of Subject D are close, and *uneune* of Subject C and *zarazara* of Subject D are close. In short, they are opposite.

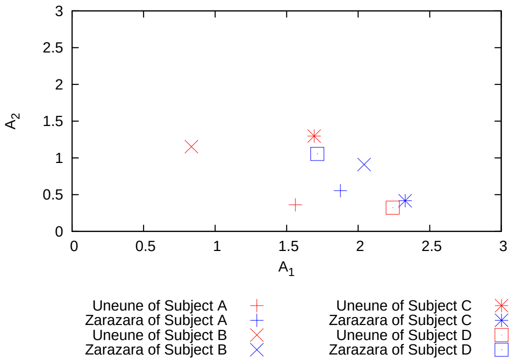


Figure 3.19: Averages of the evolved RNNs' outputs $A_1(n)$ and $A_2(n)$ for 20 seconds

We show the patterns of the RNNs' outputs $A_1(n)$ and $A_2(n)$ for 20 seconds in Figure 3.25. As we also have described the average of $A_1(n)$ and $A_2(n)$ above, the sizes of $A_1(n)$ and $A_2(n)$ of *uneune* evolved by Subject C are similar to those of *zarazara* evolved by Subject D, and the sizes of $A_1(n)$ and $A_2(n)$ of *zarazara* evolved by Subject C are similar to those of *uneune* evolved by Subject D. Namely, here they are also opposite.

Moreover, the *uneune* pattern of Subject A has relatively low frequency and high amplitude waves, and $A_2(n)$ is higher than $A_1(n)$ only in the *uneune* pattern of Subject B. Therefore, the *uneune* patterns seem to be more varied than *zarazara*.

Figure 3.26 shows the power spectrums of the RNN outputs $A_1(n)$ and $A_2(n)$ for 20 seconds, and Figure 3.27 shows the voltage outputs $V(t)$ calculated from the $A_1(n)$, $A_2(n)$, and Equation 3.1. In Figure 3.27, there are 2 peaks around 30 and 180 Hz, because f_1 and f_2 used in the equation are 30 and 180 Hz, respectively. We have described that SA I responds well under 2 Hz of vibration in Section 3.2.1. In the patterns of $A_1(n)$ and $A_2(n)$ of *uneune*, and $A_2(n)$ of *zarazara* of Subject A in Figure 3.26, the power spectrums of $A_1(n)$ and $A_2(n)$ in the low-frequency range (0.5 - 2 Hz) are high. However, in Figure 3.27, the power spectrums of $V(t)$ in the frequency range are much smaller than 30 and 180 Hz. Thus, a vibration that stimulates SA I is not found. However, because the figures for the power spectrums are generated from all ranges of the data for 20 seconds, SA I might respond to temporary stimulations that cannot be observed in the power spectrums.

3.5.3 Hand movements

We show the trajectories of the hands' positions on the XY coordinates for 20 seconds in Figure 3.28. The red points "+" represent the starting points of the hand's movements, the green lines represent the trajectories, and the blue points "x" represent the end points of the hands' movements. Both the trajectories of *uneune* and *zarazara* of Subject A are slow back-and-forth motions, both of Subject B are wandering motions, and both of Subject C are faster back-and-forth motions than Subject A. However, Subject D changes the patterns of his hand movements according to *uneune* and *zarazara*. In *uneune* of Subject D, the trajectory is a wandering motion, but in *zarazara* of Subject D, the trajectory is a back-and-forth motion.

Figure 3.29 shows the frequency distribution of the hands' velocity Ve for the 20 seconds. The interval range in the figure is 0.01. It looks that the totals of the frequency distributions are less than 1, but this is caused by the fact that the values of $Ve = 1$ are high. Both distributions of *uneune* and *zarazara* of Subject A are almost uniform. Both of Subject B are convex. Both of Subject C are also almost uniform, but they are lower than Subject A. The distribution of *uneune* of Subject D is convex, but the distribution of *zarazara* of Subject D is almost uniform, and the value of $Ve = 1$ is high. Namely, Subject D changes Ve according to the onomatopoeias.

Figure 3.30 shows the frequency distribution of the hands' acceleration Ac for 20 seconds. The interval range in the figure is also 0.01. Here also, the distributions of *uneune* and *zarazara* of Subjects A, B, and C are similar to each of the onomatopoeias. However, the distributions of *uneune* and *zarazara* evolved by Subject D are different from each other between the sensations. In the distribution of *uneune* of Subject D, the values in the low area of Ac are high, and in that of *zarazara*, the values are almost uniform. This result also suggests Subject D changes it according to the onomatopoeias.

Figure 3.31 shows the power spectrums of Ve and Ac for 20 seconds. Both Ve power spectrums of *uneune* and *zarazara* of Subject A are high in a range from 0.5 to 1.0 Hz. We consider that the power spectrums indicate periodicity of slow back-and-forth motions as shown in Figure 3.28. In the power spectrums of Ve of *uneune* and *zarazara* of Subject C and *zarazara* of Subject D, the power spectrums around 2 Hz are high. The power spectrums indicate periodicity of fast back-and-forth motions as shown in Figure 3.28. However, in Figure 3.28, the trajectories of *uneune* and *zarazara* of Subject B and the trajectories of *uneune* of Subject D show wandering motions, and there is not a clear peak in the power spectrums in Figure 3.31.

3.5.4 Distinction tests

3.5.4.1 Experiment 2

The results of Experiment 2 show that the subjects distinguished between their own evolved sensations of *uneune* and *zarazara* in Table 3.4. The values not given in the parentheses in the table are the numbers of correct answers of the distinction tests, and the values given in the parentheses are evaluations of question (2) of the questionnaires as described in Section 3.4.5.

The numbers of correct answers by Subjects A, B, and C are 15 and above in 20 trials. Namely, they are in the significance level as described in Section 3.4.3. Therefore, we consider that Subjects A, B, and C recognized the differences between their evolved sensations of *uneune* and *zarazara*. However, Subject D's number of correct answers is less than 15. We wondered at Subject D's result, because the A_1 and A_2 sizes of *uneune* and *zarazara* evolved by Subject D are too different from each other between the onomatopoeias as shown in Figure 3.25. Therefore, we conducted the test with Subject D again, and then Subject D answered all the trials of the distinction test correctly. We think that the difference between Subjects A, B, and C and Subject D is caused by their hand movements. Subjects A, B, and C used almost the same hand movements between *uneune* and *zarazara* to evolve the sensations as described in Section 3.5.3. Subject D, however, changed his hand movements according to the sensations. Therefore, this distinction test seems to be difficult for Subject D.

Subject C, who answered perfectly, evaluated 1 as "strongly agree" in the questionnaires. Subjects A and B, whose number of correct answers is less than 20 but within the significance level, evaluated 3 as "undecided." However, Subject D, whose number of correct answers is not within the significance level, evaluated 2 as "agree." Namely, the evaluations of Subjects A, B, and C's questionnaires almost fit each the number of correct answers, but Subject D's do not fit it.

Table 3.4: Results of the distinction tests between the subjects' own evolved sensations of *uneune* and *zarazara*

Subject	*uneune* vs *zarazara*
A	15(3)
B	17(3)
C	20(1)
D	12(2)

3.5.4.2 Experiment 3

Table 3.5 shows the results of Experiment 3 in which the subjects distinguished between the sensations with their own normal movement and the dummy movement. The values not given in the parentheses in the table are the numbers of correct answers for Experiment 3. The left values given in the parentheses are evaluations of question (2) in the questionnaires, and the right values are evaluations of question (3). The number of correct answers for *uneune* for Subject A is only 5. We conclude that Subject A misunderstood the 2 sensations, but distinguished between the sensations as described in Section 3.4.3. However, Subjects B, C, and D did not distinguish between the sensations. The tests seems to be difficult. The evaluations of question (2) except *uneune* of Subjects A and C also indicate the difficulty of the tests. The evaluations of question (3) almost follow those of question (2).

Table 3.5: Results of the distinction tests between the subjects' own normal movement and the dummy movement

Subject \ Onomatopoeia	*uneune*	*zarazara*
A	5(2,2)	8(4,4)
B	12(4,5)	14(4,5)
C	10(2,2)	7(4,4)
D	9(4,3)	10(4,3)

3.5.4.3 Experiment 4

The results of Experiment 4 are shown in Tables 3.6 and 3.7. Table 3.6 shows the *uneune* results, and Table 3.7 shows the *zarazara* results. The values not given in the parentheses in the tables are the numbers of correct answers in Experiment 4, and the values in the parentheses are the evaluations of question (2) as described in Section 3.4.5. "Target" means a person who evolves another sensation that is not the sensation evolved by the subject who is doing the test then. For example, in Table 3.6, the right upper number 16(2) means the results that Subject A distinguished between the sensations he evolved and those evolved by Subject D.

For *uneune*, the number of tests in which the sensations could not be distinguished between is 3 as counted in Table 3.6. For *zarazara*, the number is 6 as counted in Table 3.7. Figure 3.20 shows the percentages of the correct answers of each subject. In the figure, when a number X of the correct answers is less than half of the trial times "20", 20 - X is used to calculate the percentages. In the case of Subjects A, B, and C, it is more difficult to make a distinction between the subject's sensation and others in the case of *zarazara* than in the case of *uneune*. Therefore, we conclude that distinguishing *zarazara* is more difficult than *uneune*.

Table 3.6: Results of the distinction tests between the subjects' own and others' *uneune*

Subject \ Target	A	B	C	D
A	\	20(1)	20(1)	16(2)
B	20(1)	\	17(2)	20(1)
C	20(1)	14(5)	\	20(1)
D	14(1)	20(1)	12(2)	\

Table 3.7: Results of the distinction test between the subjects' own and others' *zarazara*

Subject \ Target	A	B	C	D
A	\	14(2)	16(2)	19(2)
B	18(2)	\	20(1)	13(4)
C	9(4)	20(1)	\	8(5)
D	2(2)	9(4)	19(1)	\

The sizes of $A_1(n)$ and $A_2(n)$ of *uneune* and *zarazara* evolved by Subjects C and D were opposite as described in Sections 3.5.2 and 3.5.3. However, in Table 3.6, Subject D could not distinguish between *uneune* evolved by Subjects C and D, and in Table 3.7, Subject C did not distinguish between *zarazara* evolved by Subjects C and D. Figure 3.21 shows the patterns of $A_1(n)$ and $A_2(n)$ of *uneune* evolved by Subject C touched by Subject D in Experiment 4, and Figure 3.22 shows the patterns of *zarazara* evolved by Subject D touched by Subject C. The patterns in Figure 3.21 are finished around 18 seconds, because

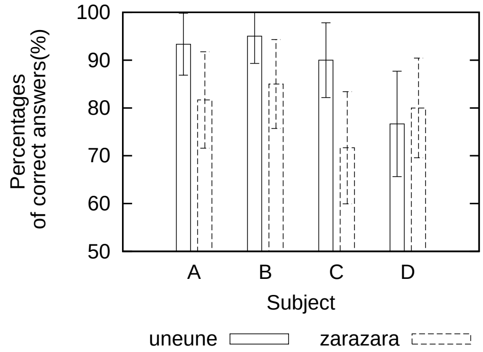
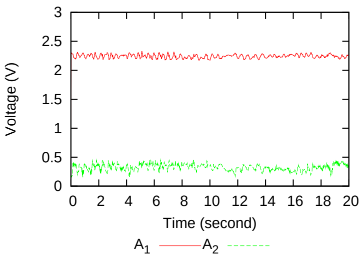


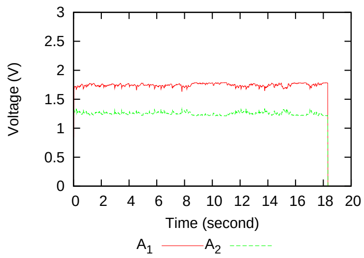
Figure 3.20: Percentages of correct answers of each subject. The error bars mean 95% confidence intervals calculated from each of the 60 trials, which includes multiplying 20 trials by 3 other subjects.

then Subject C decided to select which sensations he touched, and stopped the system by himself. In Figure 3.21, the sizes of $A_1(n)$ and $A_2(n)$ are very different from each other between the sensations. However, Subject D could not distinguish between them. In Figure 3.22, the sizes of $A_1(n)$ and $A_2(n)$ are also very different from each other between the sensations, but Subject C could not distinguish between them.

Moreover, as shown in Table 3.7, Subject C distinguished between *zarazara* he evolved and that evolved by Subject B. Figure 3.23 shows the sizes of $A_1(n)$ and $A_2(n)$ of *zarazara* evolved by Subject B touched by Subject C in Experiment 4. The sizes of $A_1(n)$ and $A_2(n)$ of *zarazara* evolved by Subject C in Figure 3.22 are much more similar to those of *zarazara* evolved by Subject B than to those of *zarazara* evolved by Subject D. If people distinguished between the sensations by using the difference between sizes of $A_1(n)$ and $A_2(n)$, Subject C could distinguish also between the sensations of *zarazara* he evolved and that evolved by Subject D. Namely, the results indicate that people do not perceive sizes of A_1 and A_2 to distinguish sensations.

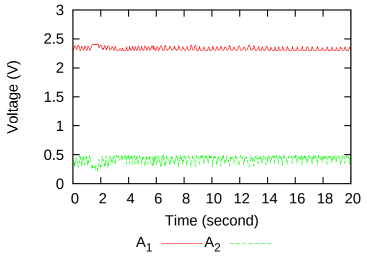


uneune evolved by Subject D touched by Subject D

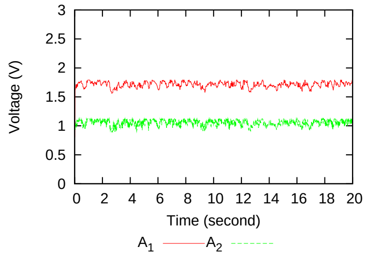


uneune evolved by Subject C touched by Subject D

Figure 3.21: Patterns of the RNN outputs $A_1(n)$ and $A_2(n)$ of *uneune* evolved by Subjects C and D touched by Subject D. The left figure is the same as the figure in Figure 3.25.



zarazara evolved by Subject C
touched by Subject C



zarazara evolved by Subject D
touched by Subject C

Figure 3.22: Patterns of the RNN outputs $A_1(n)$ and $A_2(n)$ of *zarazara* evolved by Subjects C and D touched by Subject C. The left figure is the same as the figure in Figure 3.25.

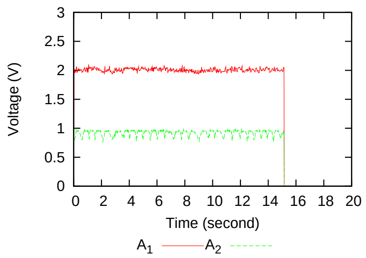


Figure 3.23: Patterns of the RNN outputs $A_1(n)$ and $A_2(n)$ of *zarazara* evolved by Subject B touched by Subject C

3.5.5 Recurrent units

We show recurrent units of the RNN for 20 seconds in Figure 3.32. In Figure 3.32, the points wander around a limited area in each figure. Figure 3.33 shows a zoomed view of the limited areas in Figure 3.32. We consider that the wanderings are internal dynamics as materials represented by onomatopoeias have, as described in Section 3.4.2.

Moreover, we inputted dummy velocity Ve and acceleration Ac in the evolved RNNs, and logged recurrent units then. The dummy Ve and Ac are given by Equation 3.8 used in Experiment 3. Figure 3.34 shows the recurrent units with the dummy Ve and Ac. Figure 3.35 shows a zoomed view of the limited areas in Figure 3.34. The areas in which points move in Figure 3.34 are not very different from those in Figure 3.32. However, in Figure 3.35, the patterns of the points do not wander but draw circles. These results indicate that hand movements are important for creating internal dynamics. On the other hand, in the results of Experiment 3, the subjects could not distinguish between sensations with their own movement and the dummy movement well. Differences between the sensations with the subjects' own movement and the dummy movement seem to be too small to be applied in the tests.

3.5.6 Noise threshold

Figure 3.24 shows results of the noise threshold for each subject in Experiment 5. The noise thresholds of *uneune* of all the subjects are less than those of *zarazara*. The results imply that *zarazara* is more robust than *uneune*.

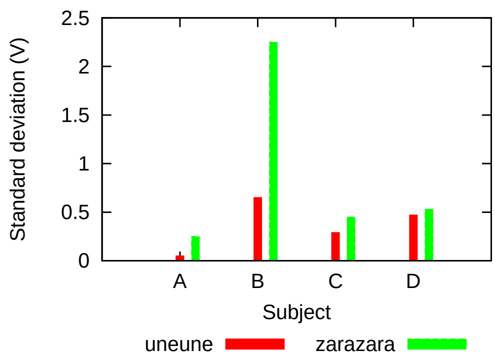


Figure 3.24: Noise thresholds as standard deviations of the normal distribution random numbers of each subject.

3.5.7 Sentences using *uneune* and *zarazara*

In Experiment 6, all *zarazara* sentences made by the subjects express the characteristics of things. Here is an example:

> (1) neko no sita wa *zarazara* si-te-iru.
> "Cat's tongue is *zarazara*."

On the other hand, most *uneune* sentences express the motions as shown in (2), or the shift of the speaker's perspective (which are called "fictive motions" in linguistics (W., 2005)) as shown in (3).

(2) nagaku-te sinayakana mono o furu-to *uneune* suru.
"Whenever we shake a long and flexible thing, it becomes *uneune*."

(3) Iroha-zaka wa *uneune* si-ta miti da.
"Iroha slope is a road, which is *uneune*."

3.6 Discussion

3.6.1 Distinctions and dynamics of the amplitude of sine waves

As we have described in Sections 3.5.2 and 3.5.4.3, the patterns of the sensations evolved by Subjects C and D are very different from each other, but Subject C did not distinguish between the sensations of *zarazara*, and Subject D did not distinguish between the sensations of *uneune*. Additionally, although the sizes of $A_1(n)$ and $A_2(n)$ of *zarazara* evolved by Subject C are more similar to those of Subject B than to those of Subject D, Subject C distinguished between *zarazara* he evolved and that evolved by Subject B. Therefore, we conclude that people do not perceive sizes of $A_1(n)$ and $A_2(n)$ to distinguish sensations.

While the evolution of the sensation in Figure 3.18 was discussed in Section 3.5.1, the averages of $A_1(n)$ and $A_2(n)$ wandered, and their movements are not similar to evolution image in Figure 3.10. This result also indicates that the sizes of $A_1(n)$ and $A_2(n)$ are not directly related to the sensation image.

Normally, to make tactile sensations, researchers control the parameters of sine waves that generate vibration of the tactile displays (e.g., (Watanabe and Fukui, 1995; Pasquero and Hayward, 2003; Konyo et al., 2000)), because it is an easy way to produce various sensations. However, our results indicate that simple combinations of sine waves cannot represent all sensations.

3.6.2 Relations between sensations and hand movements

The subjects' hand movements are different from each other between the subjects as we have described in Section 3.5.3. Thelen et al. claimed that infants acquire reaching skills from individuals' intrinsic dynamics (Thelen et al., 1993). Adults' intentional movements such as reaching are not very different from each other, but we conclude that adults also have intrinsic dynamics in their active movements to feel tactile sensations. In addition, when the subjects touched sensations evolved by others, the subjects moved their hands in their own manner. We suppose that then the subjects feel different sensations from sensations touched by the person who evolves the sensations. Subject D could not make a distinction between the sensations in the first challenge of Experiment 2. When Subject D evolved his sensations in Experiment 1, he changed his movements for *uneune* and *zarazara*. We believe that it is difficult for Subject D to decide which movement he should adopt for strange sensations. Therefore, it appears that sensations are related to hand movements. People can feel multiple sensations from one real material by changing their own hand movements. For example, people can feel unevenness and smoothness from a soft towel with slow and fast hand movements, respectively.

However, in Experiment 3, the subjects could not distinguish between the sensations with their own movement and dummy movement except in the case of *uneune* of Subject A. The results in Experiment 3 seem to contradict the relations between sensations and hand movements we described above.

We described classes of active touch in Section 1.3. If people move their hands in the way based on relations between sensations and hand movements to feel sensations, their active touch is classified under Class 3 "Input and output relationship." If people move their hands independently of sensations, their active touch is classified under Class 2 "Temporal aspects of sensory inputs." For example, if people touch sandpaper passively by guiding their fingers, they can feel the same sensation as when they touch sandpaper actively, because sandpaper does not have internal dynamics normally. Class 2 is this type. We suppose that a class of active touch of *zarazara* is Class 2, because *zarazara* means "touch of coarse surface" like sandpaper. *Uneune* means "touch of winding things," and the meaning of *uneune* depends on people more than the meaning of *zarazara*, as we will explain in Section 3.6.4. Therefore, we can consider that *uneune* of only Subject A has the meaning of Class 3 of active touch.

We have other ideas for why Subjects B, C, and D could not distinguish between *uneune* with their own movements and dummy movements in Experiment 3. One idea is that even *uneune* may not have a strong meaning of Class 3 of active touch. For example, *fuwafuwa* (fluffy) and *bunibuni* (spongy) should represent the elasticity and hysteresis of tactile materials more than *uneune* and *zarazara*. In addition to adopting such onomatopoeias, we can also add mechanisms such as a kinesthetic display to the system to represent elasticity explicitly. We will discuss the extension of the system in Section 3.6.5.

Our second idea is that it may be too difficult to distinguish between the sensations of one's own movements and dummy movements. When people touch objects actively, they can try to touch objects again and again. However, when people touch objects passively like the tactile display with dummy movements, they receive few hints with which recognize the objects. Therefore, we should conduct experiments using dummy movements for both sensations. For example, if a subject tries to make a distinction between sensations evolved by himself and others with dummy movements, and the results are worse than the results of Experiment 4, then we can discuss the contributions of hand movements more.

3.6.3 Internal dynamics of RNNs

The RNNs had internal dynamics in the recurrent units as we have described in Section 3.5.5. We believe that internal dynamics can represent states of real tactile materials such as bumps on the surface, hysteresis as a hair condition, elasticity and others, because internal dynamics can have memories as states of the recurrent units.

When the RNNs received inputs of dummy movements in Experiment 3, the recurrent units drew only a circle. However, even if the dynamics of the recurrent units is periodic, the dynamics of the outputs would not become simple combinations of sine waves. Therefore, the RNNs even with dummy movements are enough to represent sensations, which is Class 2 of active touch.

3.6.4 Differences between *uneune* and *zarazara*

In Experiment 4, in the case of Subjects A, B, and C, it is more difficult to make a distinction between the subject's sensation and others in the case of *zarazara* than in the case of *uneune*. The results show that *uneune* has less publicity than *zarazara*. We communicate with each other by words that we share, but the meanings of the words are a little different for each person, and variation ranges of the meanings are also a little different for each word. We think that the larger a variation range a word has, the less publicity the word has. Ikegami and Zlatev (Ikegami and Zlatev, 2008) emphasize that onomatopoeias typically show a close connection between body image and linguistic meaning. We believe that the more a subject uses his body in perception, the more active his perception is. When the degree of active perception is higher, the sensation owes more to the subject's bodily movement that is difficult to share with others. That is, the results imply *uneune* has a higher degree of active perception compared to *zarazara*.

In Experiment 5, the noise thresholds of *uneune* of all the subjects are less than those of *zarazara*. The results also show that *uneune* has a higher degree of active perception than *zarazara*. The subjects move their own hands to detect differences between normal sensations and sensations with noise. We consider that the subjects explore sensations to find hints to detect the differences. If the subjects find hints, they can touch the sensations in the same condition (e.g., their hand speed and position) in which they find hints again and again. Therefore, we conclude that active perception causes less robustness.

Experiments 4 and 5 have to do with the reconstruction of bodily image using linguistic inputs. In contrast, Experiment 6 has to do with linguistic usage. In Experiment 6, only *uneune* was used to describe a fictive motion. In a fictive motion sentence, while the predicate expresses dynamic motion, the sentence does not express a dynamic event. It is the speaker's perspective that moves dynamically. For example, in example (3) in Section 3.5.7, the sentence has dynamic reading, and it is not the road, but the speaker's perspective, that is moving in *uneune* motion. That is, any fictive motion sentence involves active perception at the linguistic level. And we may say that *uneune* has a higher degree of active perception than *zarazara*.

All results show that *uneune* has more intensive characteristics of active perception compared to *zarazara*. Thus, we have shown that, among onomatopoeias, some require more active perception than others. The above argument is summarized in Table 3.8.

Table 3.8: Higher degrees of active perception

	Bodily analysis		Linguistic analysis
	(Experiment 4) More private sensation	(Experiment 5) More fragile sensation	(Experiment 6) Expression utilizing speaker's perspective more
uneune	○	○	○
zarazara	×	×	×

Now the problem is how two active perceptions in body and language are connected. As we have seen in Experiment 4, the publicity of the bodily meaning of onomatopoeias is one aspect of active perception at the bodily level. What we have to note is that there is a high possibility that sharing of bodily meaning of onomatopoeias with others has some effects on sharing of their linguistic meanings with others. We assume that we might find a more direct connection between bodily and linguistic aspects of onomatopoeias by focusing on the relation between the semantic stability (i.e., how much people share the semantics of the expression) and active perception (Uno et al., 2009).

3.6.5 Improvement of the system

We conclude that our system can generate multiple tactile sensations, because Subjects A, B, and C made a distinction between their 2 sensations in Experiment 2, and Subject D also made a distinction in the second challenge. In addition, the subjects made a distinction between their own sensation and others' also in most cases in Experiment 4. We insist that the system with a tactile display and IEC for neural networks is effective for researching tactile sensations.

Moreover, we consider that the system can be improved in some ways. For example, when ICPFs are used for a long time, the ICPFs' power weakens, because the ICPFs lose water from the inside. Therefore, in the experiments, we allowed the ICPFs to rest sometimes, and we spent a long time for the experiments. If the time becomes short, we can try many other onomatopoeias. We will try other actuators such as a piezoelectric element.

As we described in Section 3.6.2, kinesthetic feedback may be necessary to represent sensations such as *fuwafuwa* (fluffy) and *bunibuni* (spongy). We are interested in such onomatopoeias, because we believe that the sensations of such onomatopoeias have richer inherent internal dynamics. Kinesthetic displays are used for studies of tactile sensation(e.g. (Sato et al., 2007)). We will check whether kinesthetic displays are useful for our research or not.

3.7 Conclusions

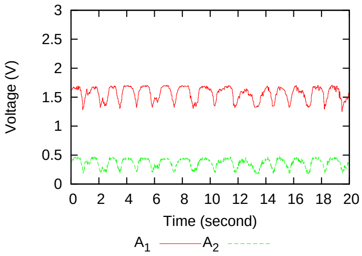
We developed a new system using a tactile display, a 3D position sensor, and RNN to research active touch, conducted experiments in which subjects evolved RNNs with IEC to emulate sensations of Japanese onomatopoeias (*uneune* and *zarazara*), and conducted tests using the evolved RNNs. The results indicated that the system can generate multiple tactile sensations, and that IEC is effective for developing a system for psychological experiments. Moreover, we discussed how to improve the system to represent more kinds of sensations.

Our results indicate that people do not perceive the amplitude of sine waves that compose tactile sensations to distinguish sensations. Therefore, we believe that a tactile display needs to be controlled by a system that is described not by simple combinations of sine waves but by algorithms such as neural networks that can have complex dynamics and hysteresis.

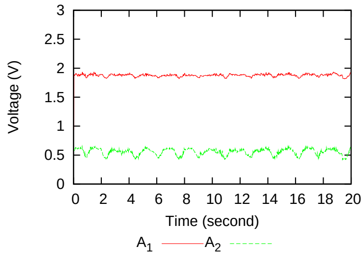
The subjects could not distinguish between sensations with normal hand movements and dummy movements well. To investigate if active touches of onomatopoeias depend on a relation between sensations and hand movements, we suggested experiments that adopt other onomatopoeias that have a meaning of elasticity and hysteresis, and a kinesthetic display that represents elasticity explicitly to the system, and that conducts a distinction test between sensations of *uneune* and *zarazara* with dummy movements.

The RNNs had internal dynamics in the recurrent units. We think that the internal dynamics represent sensations that are generated by dynamics that cannot be composed of simple combinations of sine waves, and a relation between perceptions and hand movements.

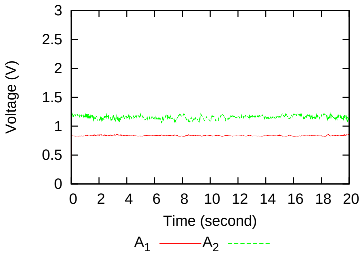
Since the subjects made a distinction between their own sensations and others' sensations of *uneune* better than those of *zarazara*, we consider that *uneune* has less publicity than *zarazara*. In addition, the results of the noise threshold test indicate that the sensations of *uneune* had the lower robustness against noise than those of *zarazara*. We did also experiments in which subjects wrote sentences using *uneune* and *zarazara*, and analyzed the semantics of onomatopoeias for tactile sensation. Then, all *zarazara* sentences made by the subjects express the characteristics of things, whereas most *uneune* sentences express the motions, or the shift in the speaker's perspective. From these results, we pointed out that *uneune* is more active than *zarazara* in perception. We also proposed that the notion of semantic stability together with active perception may become a key to finding out the connection between bodily and linguistically active perceptions.



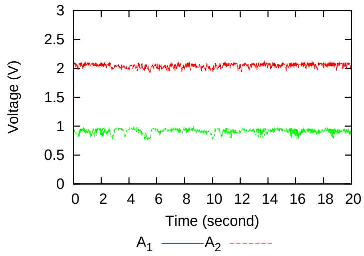
(a) *uneune* of Subject A



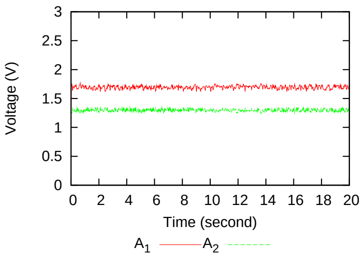
(b) *zarazara* of Subject A



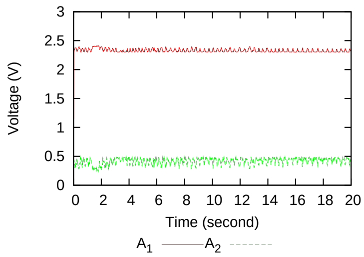
(c) *uneune* of Subject B



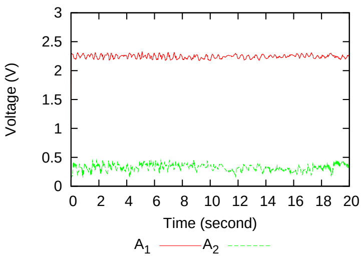
(d) *zarazara* of Subject B



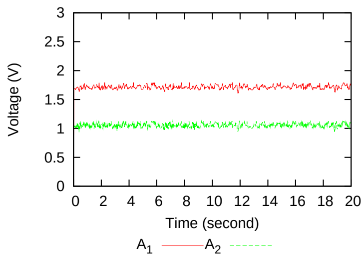
(e) *uneune* of Subject C



(f) *zarazara* of Subject C

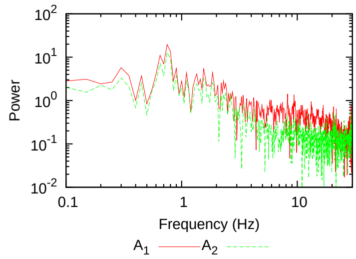


(g) *uneune* of Subject D

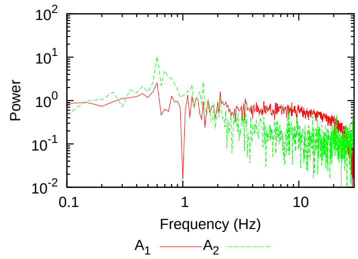


(h) *zarazara* of Subject D

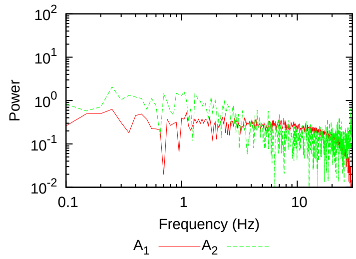
Figure 3.25: Patterns of the RNN outputs A_1 and A_2 for the 20 seconds



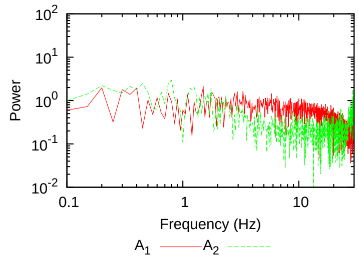
(a) *uneune* of Subject A



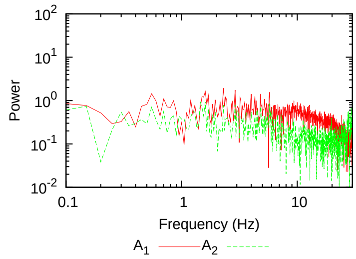
(b) *zarazara* of Subject A



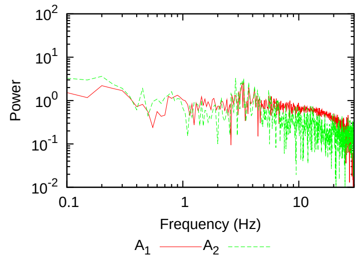
(c) *uneune* of Subject B



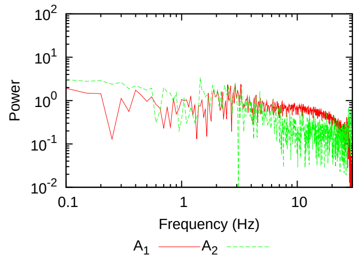
(d) *zarazara* of Subject B



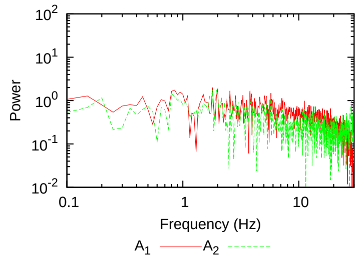
(e) *uneune* of Subject C



(f) *zarazara* of Subject C

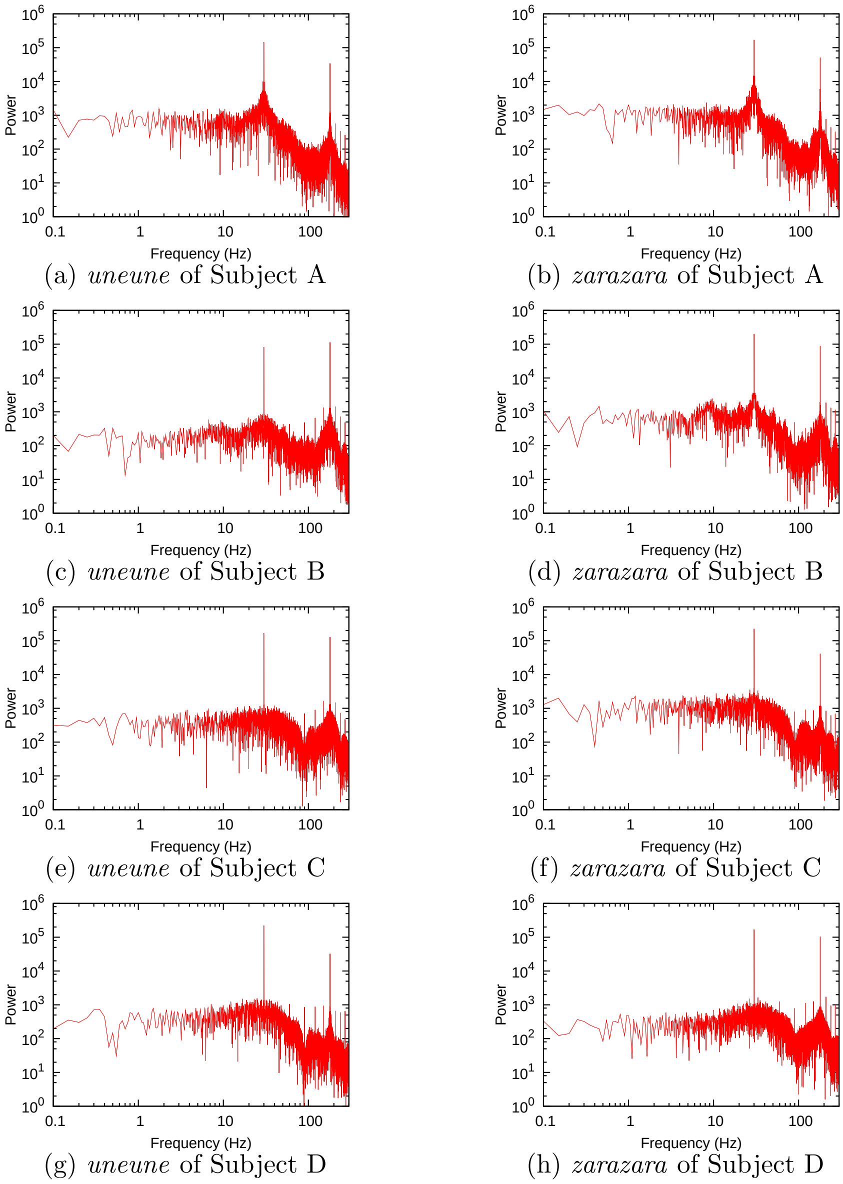


(g) *uneune* of Subject D



(h) *zarazara* of Subject D

Figure 3.26: Power spectrums of the RNN outputs A_1 and A_2 for the 20 seconds



(a) *uneune* of Subject A

(b) *zarazara* of Subject A

(c) *uneune* of Subject B

(d) *zarazara* of Subject B

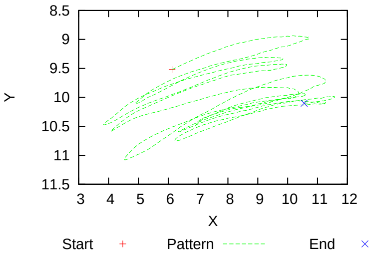
(e) *uneune* of Subject C

(f) *zarazara* of Subject C

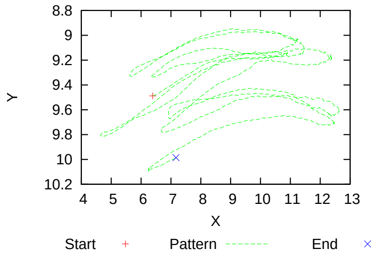
(g) *uneune* of Subject D

(h) *zarazara* of Subject D

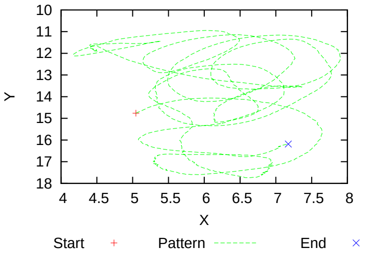
Figure 3.27: Power spectrums of the voltage outputs $V(t)$ for the 20 seconds



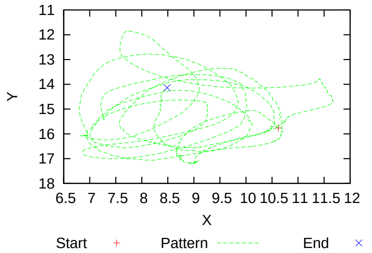
(a) *uneune* of Subject A



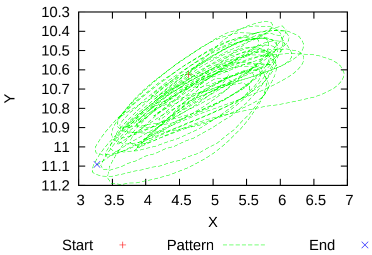
(b) *zarazara* of Subject A



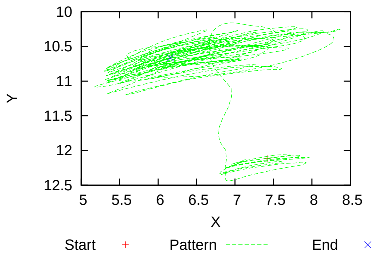
(c) *uneune* of Subject B



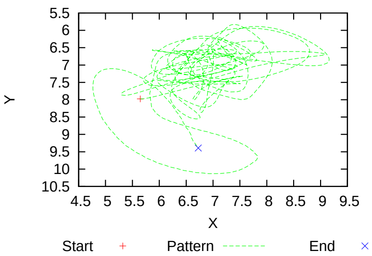
(d) *zarazara* of Subject B



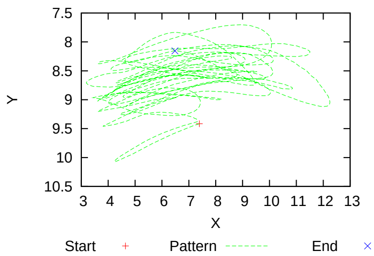
(e) *uneune* of Subject C



(f) *zarazara* of Subject C

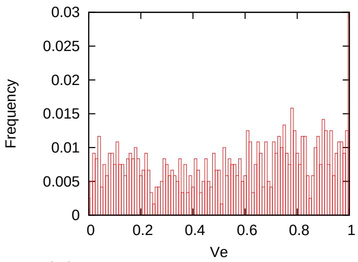


(g) *uneune* of Subject D

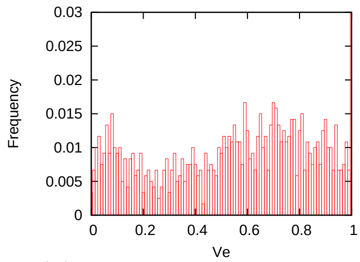


(h) *zarazara* of Subject D

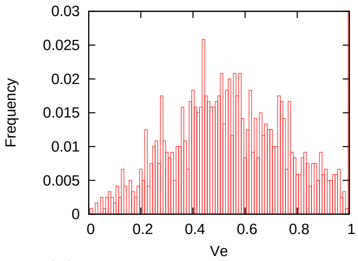
Figure 3.28: Trajectories of the hands' positions on XY coordinates for the 20 seconds



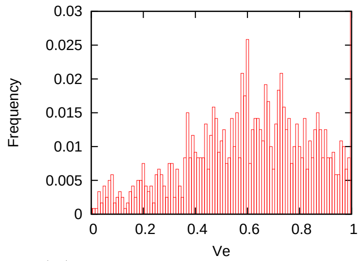
(a) *uneune* of Subject A



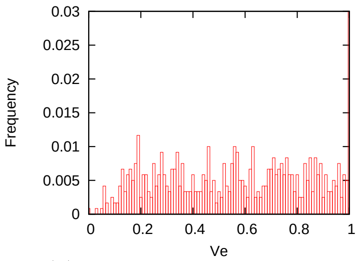
(b) *zarazara* of Subject A



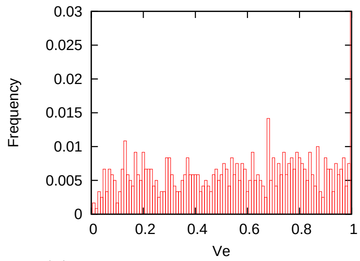
(c) *uneune* of Subject B



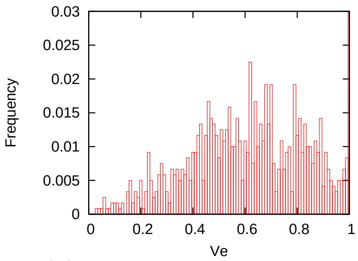
(d) *zarazara* of Subject B



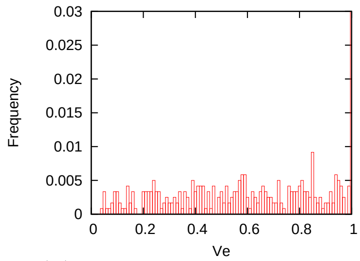
(e) *uneune* of Subject C



(f) *zarazara* of Subject C

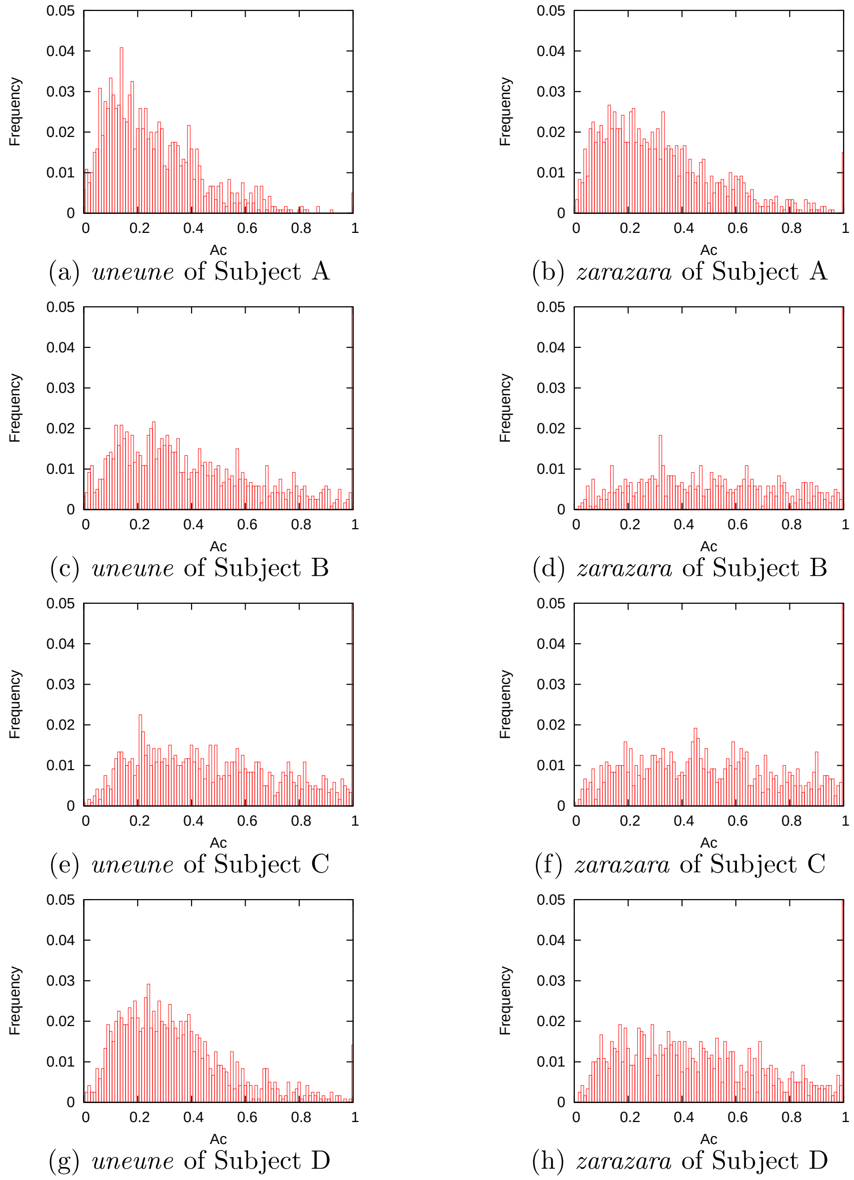


(g) *uneune* of Subject D



(h) *zarazara* of Subject D

Figure 3.29: Frequency distribution of the hands' velocity Ve for the 20 seconds



(a) *uneune* of Subject A

(b) *zarazara* of Subject A

(c) *uneune* of Subject B

(d) *zarazara* of Subject B

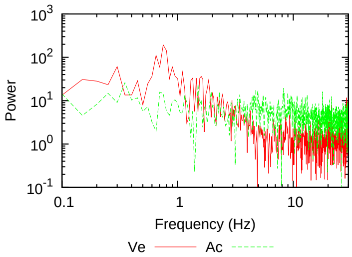
(e) *uneune* of Subject C

(f) *zarazara* of Subject C

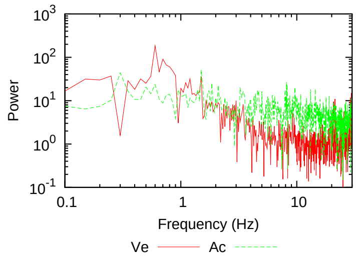
(g) *uneune* of Subject D

(h) *zarazara* of Subject D

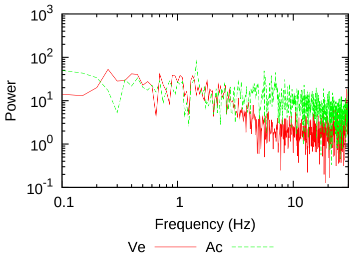
Figure 3.30: Frequency distribution of the hands' acceleration Ac for the 20 seconds



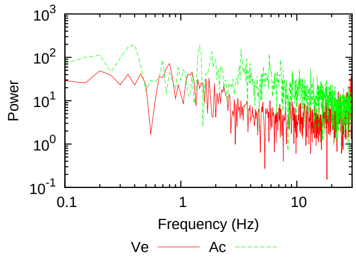
(a) *uneune* of Subject A



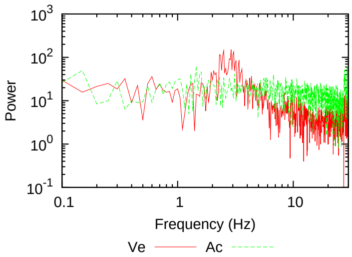
(b) *zarazara* of Subject A



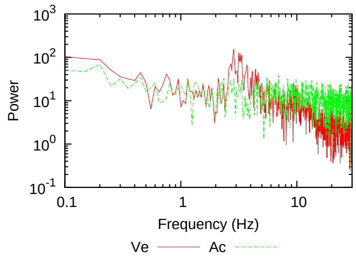
(c) *uneune* of Subject B



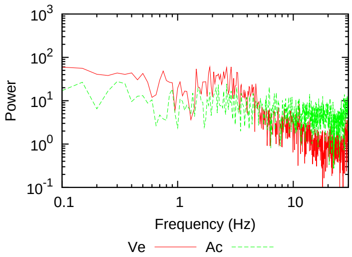
(d) *zarazara* of Subject B



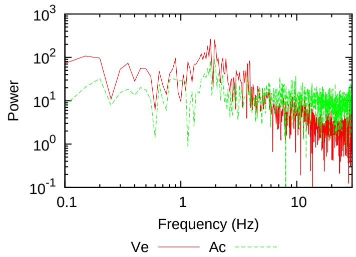
(e) *uneune* of Subject C



(f) *zarazara* of Subject C

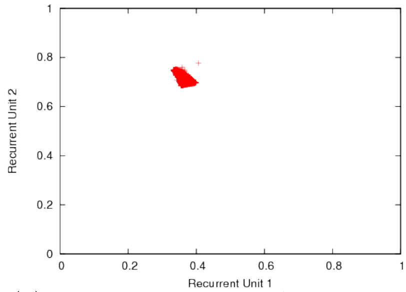


(g) *uneune* of Subject D

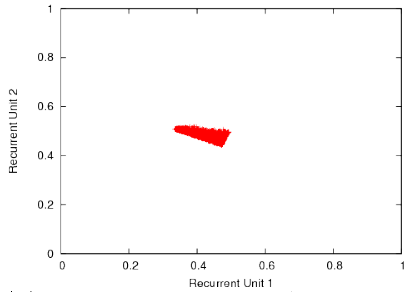


(h) *zarazara* of Subject D

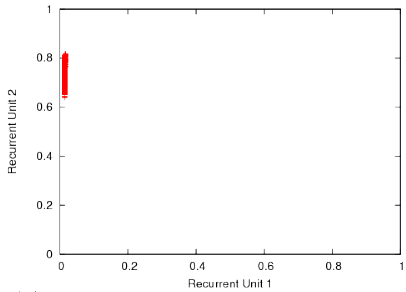
Figure 3.31: Power spectrums of the hands' velocity Ve and acceleration Ac for the 20 seconds



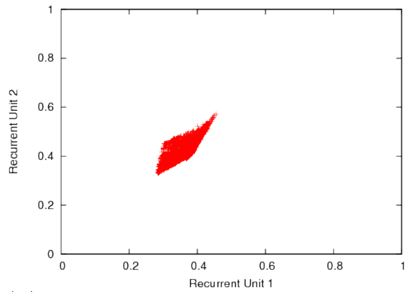
(a) *uneune* of Subject A



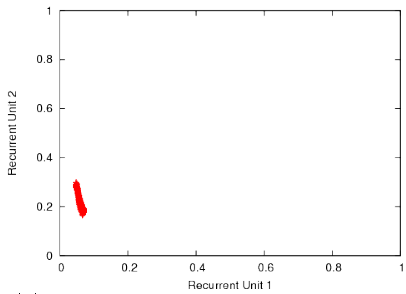
(b) *zarazara* of Subject A



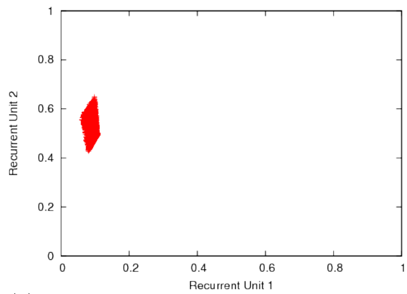
(c) *uneune* of Subject B



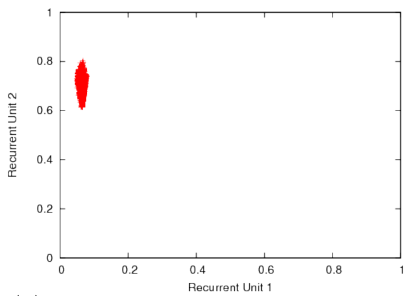
(d) *zarazara* of Subject B



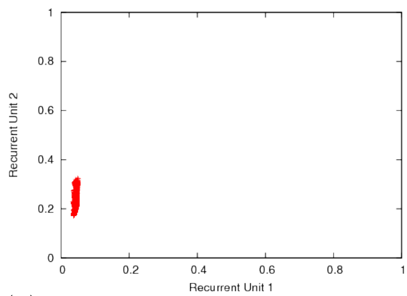
(e) *uneune* of Subject C



(f) *zarazara* of Subject C

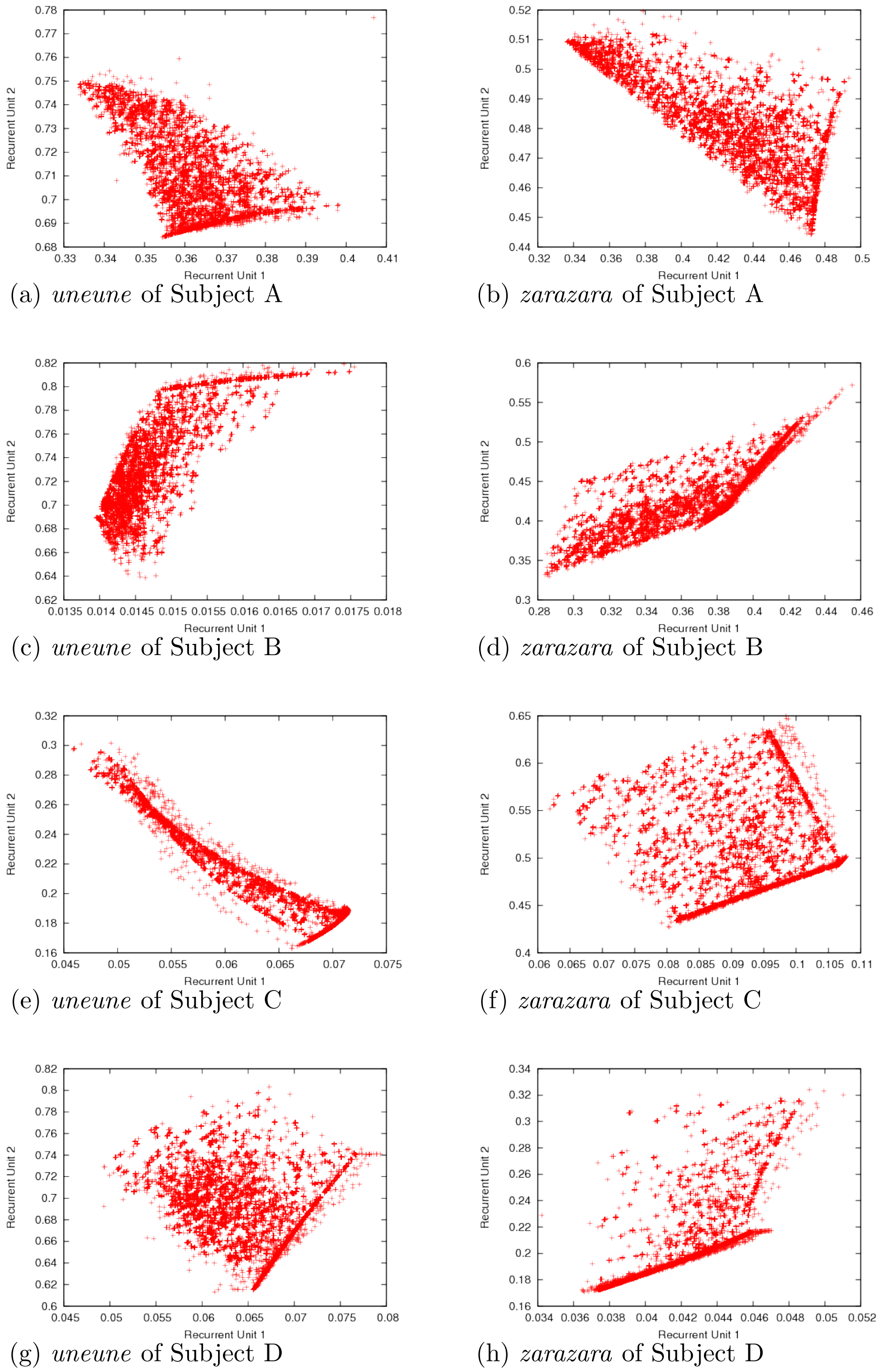


(g) *uneune* of Subject D



(h) *zarazara* of Subject D

Figure 3.32: Recurrent units of the evolved RNNs for the 20 seconds



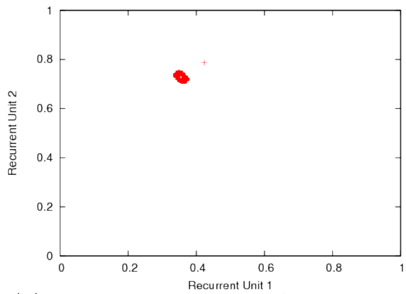
(a) *uneune* of Subject A (b) *zarazara* of Subject A

(c) *uneune* of Subject B (d) *zarazara* of Subject B

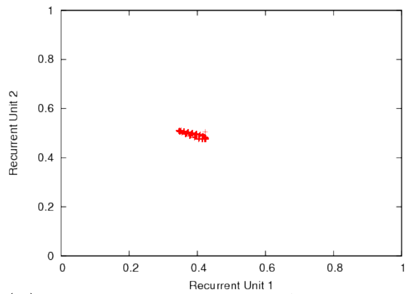
(e) *uneune* of Subject C (f) *zarazara* of Subject C

(g) *uneune* of Subject D (h) *zarazara* of Subject D

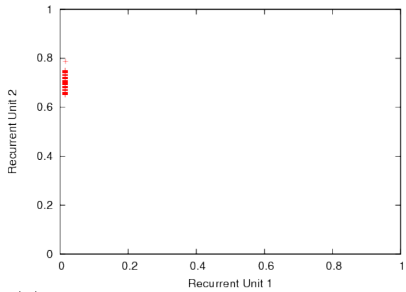
Figure 3.33: Zoom of recurrent units of the evolved RNNs for the 20 seconds



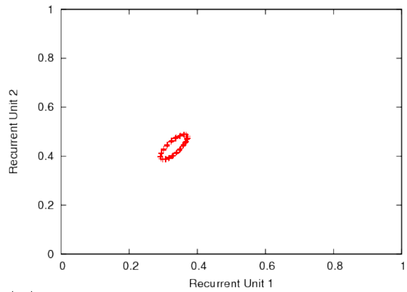
(a) *uneune* of Subject A



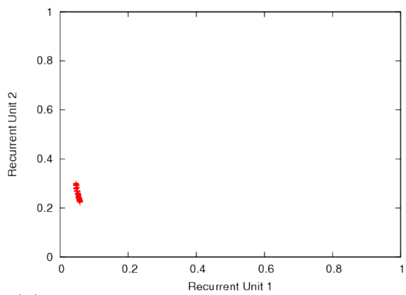
(b) *zarazara* of Subject A



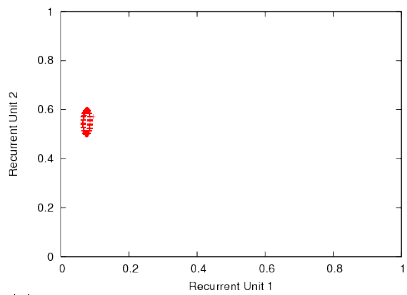
(c) *uneune* of Subject B



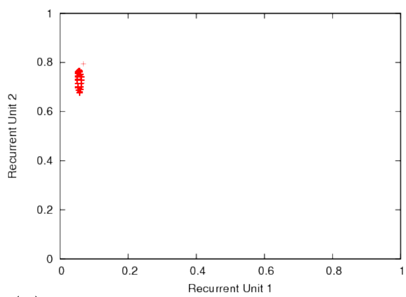
(d) *zarazara* of Subject B



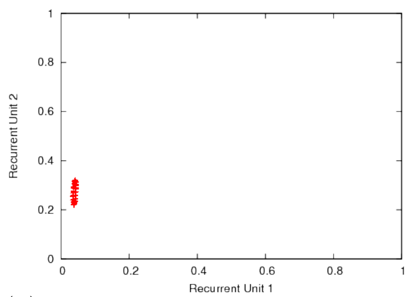
(e) *uneune* of Subject C



(f) *zarazara* of Subject C

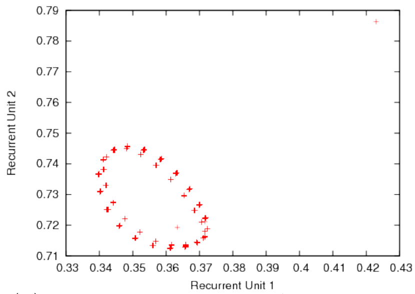


(g) *uneune* of Subject D

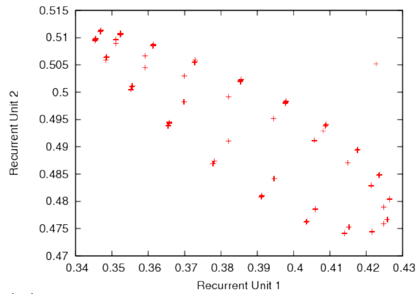


(h) *zarazara* of Subject D

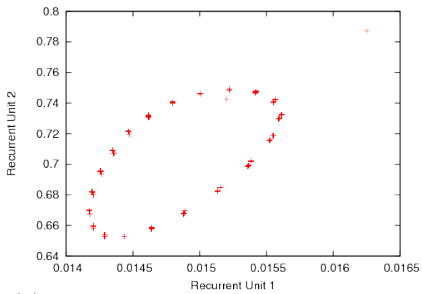
Figure 3.34: Recurrent units of the evolved RNNs with sine wave inputs



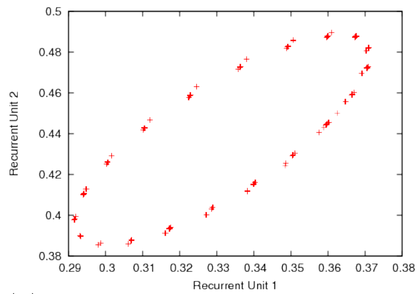
(a) *uneune* of Subject A



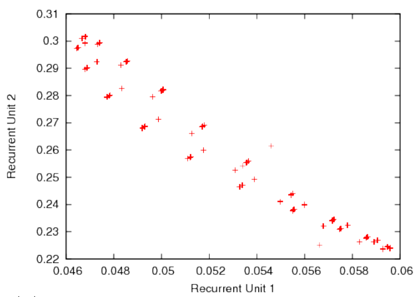
(b) *zarazara* of Subject A



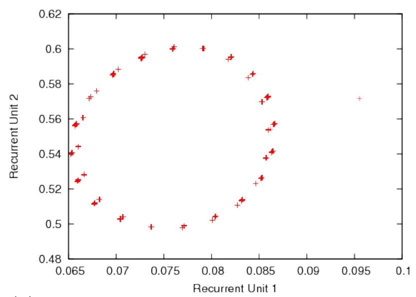
(c) *uneune* of Subject B



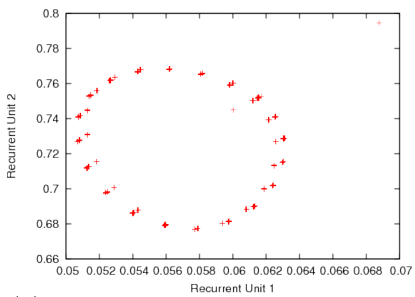
(d) *zarazara* of Subject B



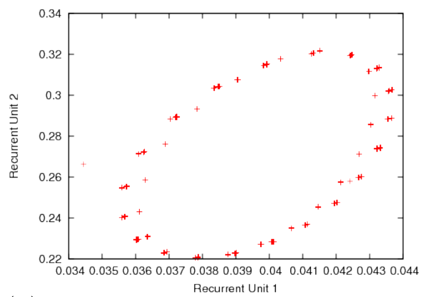
(e) *uneune* of Subject C



(f) *zarazara* of Subject C



(g) *uneune* of Subject D



(h) *zarazara* of Subject D

Figure 3.35: Zoom of recurrent units of the evolved RNNs with sine wave inputs

Chapter 4

General Discussion

4.1 Active perception and instability

The 2 studies in this thesis took the constructive approach with evolutionary neural networks (NN) and in terms of the ecological world view. The whole system is designed and controlled by deterministic agents and programs, but the environment is open and non-deterministic.

In Chapter 2, the microslip phenomena of the evolved agent can be attributed to a riddled basin-like structure. The spatial layouts of objects generate complicated reaching behavior and basins of selection. We consider that the agents have inherent structures that enhance the slight variations in spatial layout by actively moving.

In Chapter 3, as we discussed in Section 3.6.4, since the subjects made a better distinction between their own sensations and others' sensations of *uneune* compared to those of *zarazara*, we consider that *uneune* has less publicity than *zarazara*. When the degree of active perception is higher, the sensation owes more to a subject's bodily movement, which is difficult to share with others. That is, the results imply *uneune* has a higher degree of active perception compared to *zarazara*. In addition, the subjects wrote more fictive motion sentences of *uneune* than *zarazara*. In a fictive motion sentence, while the predicate expresses dynamic motion, the sentence does not express a dynamic event. It is the speaker's perspective that moves dynamically. Therefore, the results also imply that *uneune* has a higher degree of active perception than *zarazara*. Moreover, the sensations of *uneune* had lower robustness against noise than those of *zarazara*. We think that the subjects explore sensations to find hints to detect differences between normal sensations and sensations with noise. Therefore, we believe that the active perception enhanced the noise, and caused the instability of *uneune*.

The results of the 2 studies indicate that active perception is related to the instabilities between the agents/subjects and the environment. We consider active perception to be the interfaces between the agents/subjects and open environments. The interfaces become unstable by relating to the environment. If the unstable interfaces receive inputs, the interfaces send outputs that depend significantly on a slight variance in the inputs to the agents/subjects.

4.2 Instability of NNs evolved in the real world

In Chapter 2, we described the behavior of the evolved NN by varying the x-coordinates of the 2 objects little by little, and found that the selection of the objects has the riddle basin-like structure as shown in Figure 2.8. However, in experiments using real humans, we cannot check the behaviors of the humans under such conditions. To draw Figure 2.8, we ran the simulation 100,000,000 times under slightly different conditions. It is impossible to use real humans for so many precise experiments. Therefore, in Chapter 3, we could not investigate attributes of the evolved NNs as dynamical systems well.

To overcome the impossibility, we suggest simulating environments for the NNs in Chapter 3. We used the dummy movements for the NNs in Experiment 3 of Chapter 3. For example, we can investigate the behaviors of the NNs by adding the dummy movements a number of times with a slightly different frequency or amplitude of the dummy movements. Then, we may find instability against a slight variation in the frequency.

Moreover, we can also try to evolve NNs for the relationship between the subjects' hand movements and outputs of the tactile display. Namely, the NNs emulate real humans. By using evolutionary data in Experiment 1 of Chapter 3, we can use a lot of data to evolve the NNs. Then, there is no useful parameters such as the frequency and amplitude of the dummy movements. However, we can add noise to the inputs and outputs of the NNs of tactile sensation as used in Experiment 5 in Chapter 3.

Thus, we can investigate the instability of NNs evolved in real environments. We expect the NNs of *uneune* are more unstable against slight variations than those of *zarazara*, because we consider that *uneune* has a higher degree of active perception compared to *zarazara*. This will be future work.

4.3 Applications in an open environment

The interfaces of active perception are adaptable to most inputs from an open environment, and sometimes behave unpredictably. We applied the idea of the interfaces between the agents and an open environment to media art works, because we believe that the interfaces' adaptability and unpredictability are useful for media art, and expect to find hints for understanding active perception more.

We programmed a robot that dances autonomously by listening to music with an NN (Aucouturier et al., 2008a; Aucouturier et al., 2008b). The robot is based on "miuro" manufactured by ZMP Inc, and has an NN that receives music as input and produces motor outputs. Music is an open environment for the robot, because the robot does not know what music the robot will listen to, and how the music progresses. The NN of the robot was not evolved, but chosen to be an interface as it has adaptability and unpredictability by hand tuning. The robot can adapt its own dance to music, and the dance is sometimes unpredictable.

Moreover, we made a visual installation "Mind Time Machine"(MTM) (Ikegami and Ogai, 2010). MTM consists of 3 screens (right, left, and above) and 15 cameras. The 15 cameras' images are decomposed into frames, and NNs control other macro processes that combine, reverse, and superpose them to make new frames. Here also, the NNs were not evolved, but chosen by hand tuning. We collected and analyzed the MTM data, and found that the dynamics of the NNs sometimes behave chaotically, and sometimes periodically. Open environment for MTM is the visual inputs of cameras, including humans' behaviors and brightness around MTM.

We made also sound software that generates sound from dynamical systems such as cellular automaton, logistic map, and tape and machine networks (Ogai et al., 2007b; Ogai et al., 2007a). The software translates 1 number into acoustic pressure. Users control the parameters of the dynamical systems in the software. The dynamical systems are not an NN, but they also are unstable due to their chaotic dynamics. The chaotic dynamics can generate sound that cannot be represented by a simple combination of sine waves. Here also, for the dynamics in the software, the open environment is a human.

The agents of these works behave adaptively and unpredictably in an open environment such as music and humans. On the other hand, for the humans, the agents of these works are an open environment. The agents and the humans actively perceive each other. We consider that the agents and the humans are exchangeable with each other. There may be differences between the agents and the humans, but we consider that evolutionary methods can handle the differences. In the next section, we suggest new evolutionary methods through the idea about exchangeability.

4.4 Evolution in multiple situations

We suggest integrating methods in which NNs are evolved with simulations and experiments. NNs are multipurpose, and used in various fields. The NNs used in our research are also similar to each other as shown in Figures 2.2 and 3.8. The NNs are exchangeable with each other by changing their structure somewhat. We consider if an NN is adaptive to an open environment, then the NN can easily adapt to another open environment.

We suppose that real humans obtain skills to adapt the various open environments they encounter, and humans combine the obtained skills to generate whole behavior, such as the agent's selection that has a riddled basin-like structure in the microslip research. Here, we suggest applying more varied kinds of environments to evolve NNs.

Normally, evolution in a simulation starts from a random state in one situation, because researchers want to estimate the process and result of the evolution objectively. However, we think that there are behaviors that cannot emerge from such random states. In Chapter 2, we used 3 tasks. In 2 of them, an agent has to get each object, and in 1 task, an agent has to get either object. Then, the agents in the research obtained behaviors that cannot be led by evolution in a single task.

We suggest 3 methods for evolving of NNs in multiple situations.

4.4.1 Multiple people in one environment

In Chapter 3, each person evolved NNs with IEC. Then, their evolutionary speeds were slow, and they spent a number of hours for each evolution. If they evolve common NNs together, their evolutionary speed becomes faster. Then, we cannot conduct experiments using sensations that are evolved only by a subject self, but conduct experiments using many sensations such as comparing onomatopoeias. For example, we were able to discuss semantic stability from the variance in the evolved NNs.

This method is applicable to other systems and sensations. For example, we can apply NNs to our sound software, and evolve them with IEC with many people. Normally, people do not have a tactile display in their houses, but have speakers for their PC. Therefore, we can distribute the sound software over the Internet, and evolve NNs shared by users. The use of crossovers of NNs for evolution with the participation of many users and NNs also becomes possible. Namely, we can evolve NNs with collective intelligence.

4.4.2 Integration of simulation and IEC

As we described above, the NNs used in our simulation and IEC research are similar to each other, and exchangeable with each other by changing their structure somewhat. Therefore, we can also evolve NNs with simulation and IEC alternately. If we can describe a fitness function of a behavior that we want, the evolutionary increases. It is better that people set a fitness function that an NN has instability in adapting to, because we consider that there are better NNs around such an NN as has instability. For example, if people touch a sensation of the RNN evolved in the microslip research, they may feel an amazing sensation that cannot be evolved only in IEC.

4.4.3 Multi-environments

We suggest applying multi-environments to evolution. We conducted a tactile sensation experiment using IEC and NNs. The NNs applied the same method for visual and auditory sensation. As we described above, we can apply NNs to our sound software, and evolve them with IEC. In the same way, we can apply IEC to the miuro and MTM NNs. For example, in miuro, people evolve miuro behaviors by observing the behaviors of some NNs, and choosing a better NN.

We consider if an NN is too unstable to adapt to a sensation, the NN can adapt to other sensations easily. When we changed the output data of the NN evolved in the tactile sensation experiment into a sound, the sound was an interesting low sound that it is normally difficult to generate in our sound software. This example is not about NNs themselves, but about NN output data. However, it indicates that we may be able to discuss similarities between sensations, because the time scales of tactile and auditory sensations correspond there. The range of audible frequencies is between about 20 and 20,000 Hz, and the range of frequencies that mechanoreceptors of tactile sensation can detect is less than 300 Hz. In spite of the difference between the ranges, people process both sensations at the same time, and the sensations overlap. For example, *zarazara* is a Japanese onomatopoeia that represents tactile sensation, but sometimes *zarazara* is used to represent sounds. For example, if people feel tactile and auditory sensations of some NNs, and answer similarity regarding each sensation, then we can discuss multimodality with this method.

Acknowledgments

I would like to acknowledge the continuing guidance and encouragement of Prof. Takashi Ikegami.

Chapter 2 is partially supported by The 21st Century COE (Center of Excellence) program (Research Center for Integrated Science) of the Ministry of Education, Culture, Sports, Science, and Technology, Japan, and the EC Agent project, sponsored by the Future and Emerging Technologies program of the European Community (IST-1940).

In Chapter 3, the author is also grateful to Prof. Naoki Saiwaki for lending a device, to Prof. Soutaro Shimada for helpful discussions, and to Prof. Ryoko Uno for her valuable cooperation. This chapter is partially supported by a grant of DENSO CORPORATION Research Laboratories, and by a Grant-in-Aid for Scientific Research on Priority Areas "Emergence of Adaptive Motor Function through Interaction between Body, Brain and Environment" from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

In Chapter 4, the author is also grateful to Prof. Jean-Julien Aucouturier and ZMP INC. for their valuable cooperation in the music robot "miuro", and to Keiichiro Shibuya and Hiroto Ebara for valuable cooperation in the sound software. This chapter is partially supported by IPA Exploratory Software Project, and by a Grant-in-Aid for challenging Exploratory Research.

Thanks are due to all my colleagues in Ikegami Laboratory with whom I have discussed this problem.

Finally, I must thank all my friends and family for their constant encouragement for the long time of seven years.

Appendix A

Here we show diagrams demonstrating the object selection in task 2, drawn as in Figure 2.8 in Chapter 2.

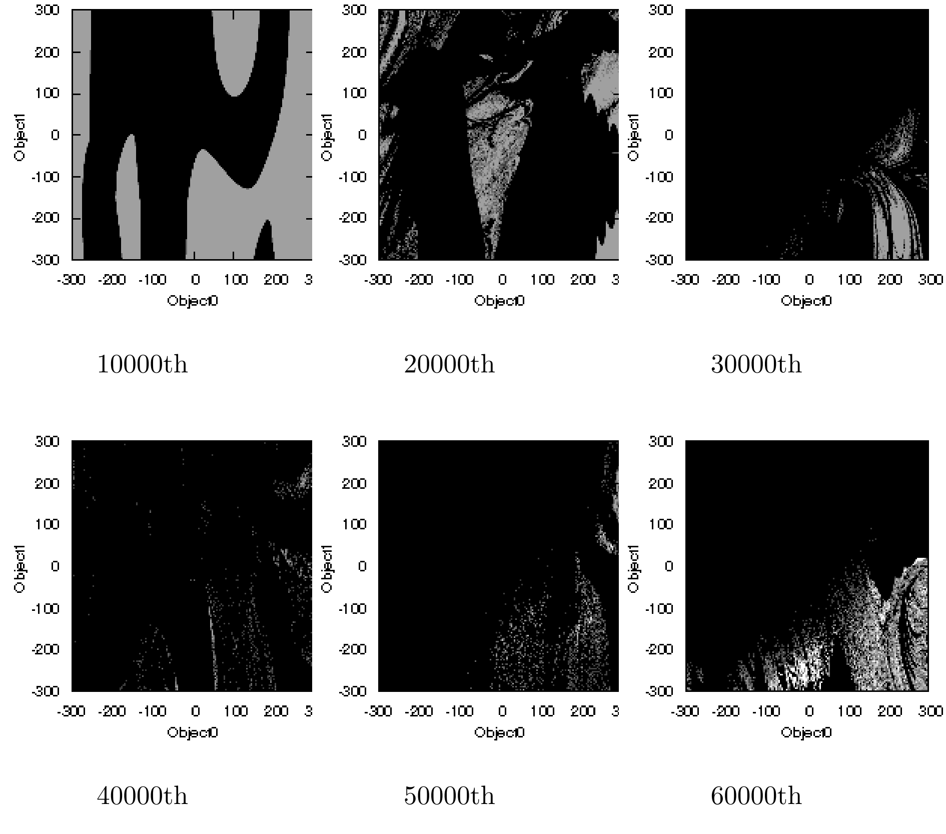


Figure 4.1: Selections of objects in task 2 by agents of 10000th - 60000th generations.

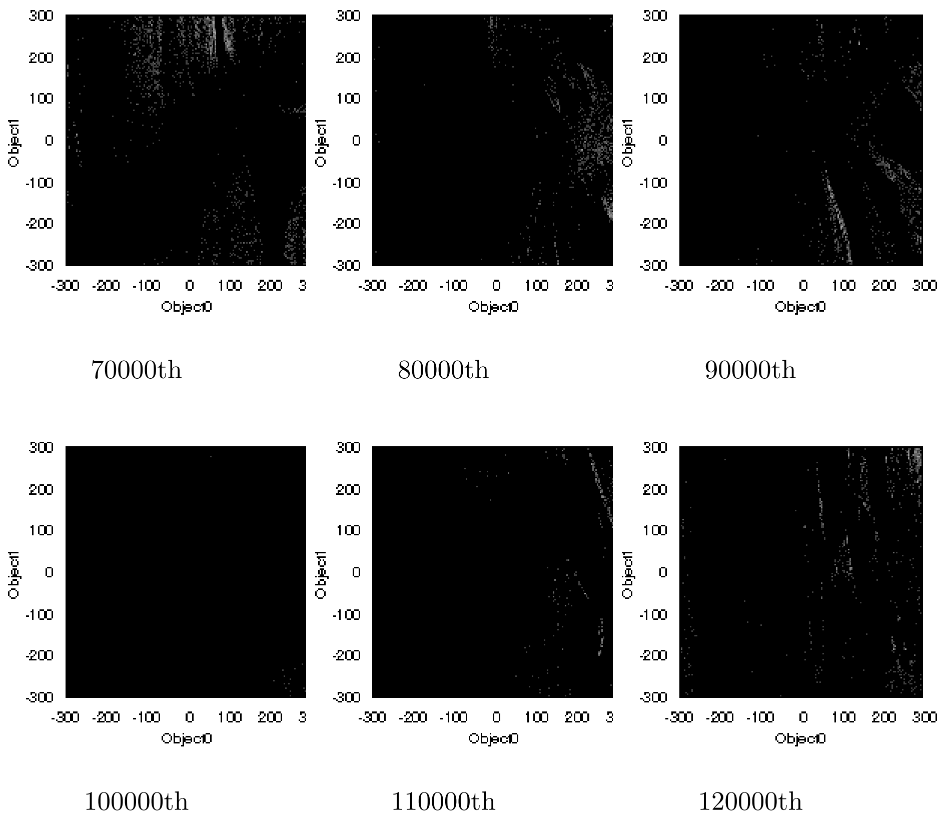
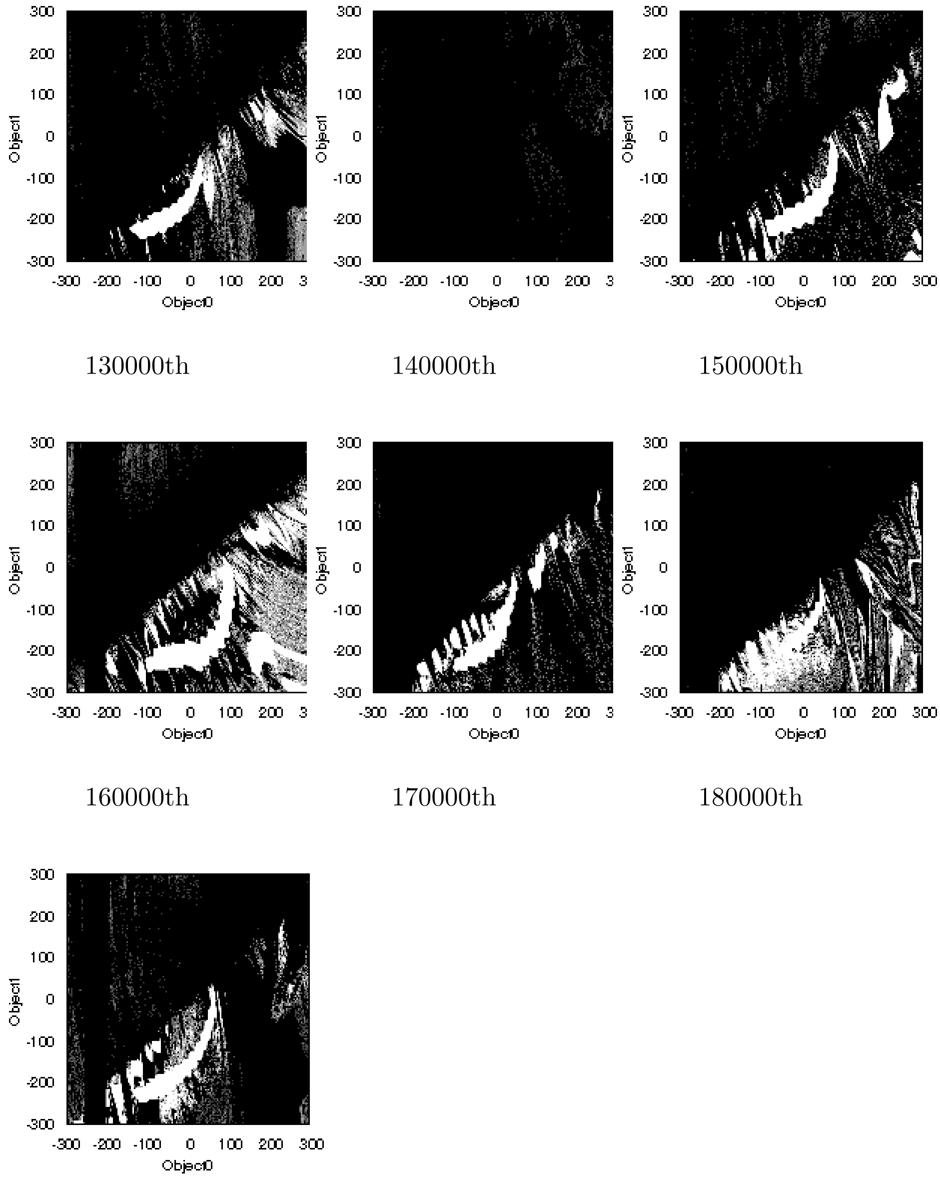


Figure 4.2: Selections of objects in task 2 by agents of 70000th - 120000th generations.



130000th 140000th 150000th

160000th 170000th 180000th

190000th

Figure 4.3: Selections of objects in task 2 by agents of 130000th - 190000th generations.

Appendix B

An another example of agent evolved with the different random seed in Chapter 2. In Figure 4.4, evolution of the fitness value is evaluated. In Figure 4.5 and 4.6, an agent of the 50000th generation is taken as an example to show the object selection pattern. The complex basin structure isn't typical but is sometimes observed in this task setup.

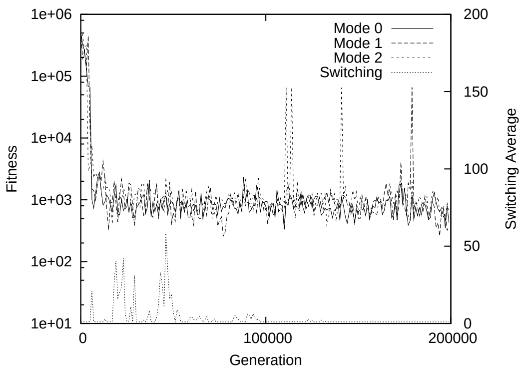


Figure 4.4: Evolution of the fitness value of each task and the number of action switching events in average (the lower line). This is drawn as Figure 2.3.

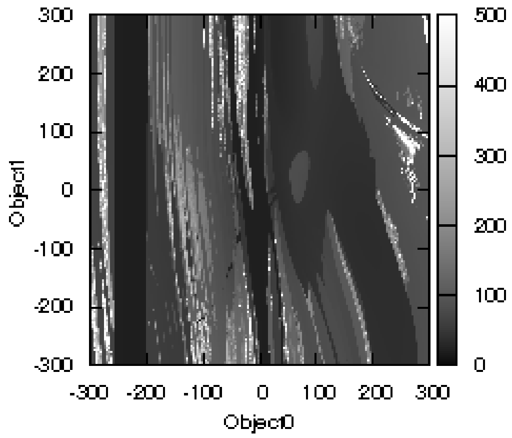


Figure 4.5: A grayscale plot of time steps required to reach an object in task 2. The darker area indicates more time steps are required to get an object. The x coordinates of object0 and object 1 are taken as a horizontal and vertical axis, respectively. This is drawn as Figure 2.6.

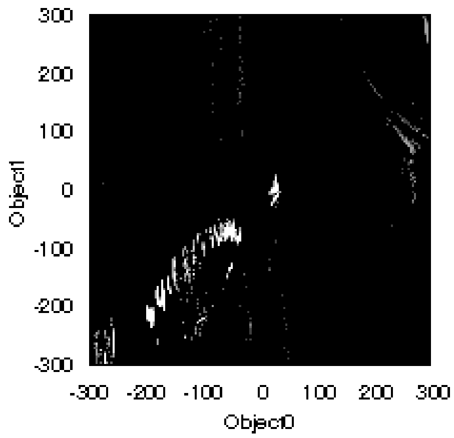


Figure 4.6: Selections of objects in task 2 by agent of 50000th generations. This is drawn as Figure 2.8.

Bibliography

Alexander, J. C., Yorke, J. A., You, Z., and Kana, I. (1992). Riddled basins. *International Journal of Bifurcation and Chaos*, 2(4):795–813.

Arons, B. (1992). A review of the cocktail party effect. *Journal of the American Voice I/O Society*, 12:35–50.

Aucouturier, J.-J., Ogai, Y., and Ikegami, T. (2008a). Making a robot dance to music using chaotic itinerancy in a network of fitzhugh-nagumo neurons. In *Neural Information Processing (Lecture Notes in Computer Science)*, volume 4985, pages 647–656. Springer Berlin/Heidelberg. (14th International Conference on Neural Information Processing (ICONIP), Kitakyushu, Japan, Nov. 2007).

Aucouturier, J.-J., Ogai, Y., and Ikegami, T. (2008b). Using chaos to trade-off synchronization and autonomy in a dancing robot. *IEEE Intelligent Systems*, 23(2). in (Aucouturier, J.-J., ed.) Trends and Controversies.

Blake, D. T., Hsiao, S. S., and Johnson, K. O. (1997). Neural coding mechanisms in tactile pattern recognition: The relative contributions of slowly and rapidly adapting mechanoreceptors to perceived roughness. *The Journal of Neuroscience*, 17(19):7480–7489.

Blakemore, S. J., Wolpert, D. M., and Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, 1(7):635–640.

Bolanowski Jr., S. J., Gescheider, G. A., Verrillo, R. T., and Checkosky, C. M. (1988). Four channels mediate the mechanical aspects of touch. *Journal of Acoustical Society of America*, 84:1680–1694.

Brooks, R. A. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, RA-2(1):14–23.

Brooks, R. A. (1991a). Intelligence without reason. pages 569–595. Morgan Kaufmann.

Brooks, R. A. (1991b). Intelligence without representation. *Artificial Intelligence Journal*, 47:139 159.

Cangelosi, A., Parisi, D., and Nolfi, S. (1994). Cell division and migration in a 'genotype' for neural networks. *Network*, 5:497–515.

Cherry, C. E. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, 25(5):975–979.

Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again*. The MIT Press.

Cornsweet, T. N. (1962). The staircase method in psychophysics. *The American journal of psychology*, 72:485–491.

Cox, K. (1964). The responsiveness of food sales to shelf space changes in supermarkets. *Journal of Marketing Research*, 1:63–67.

Darwin, C. (1881). *The Formation of Vegetable Mould, Through the Action of Worms, With Observations on Their Habits*. John Murray, London.

Dawkins, R. (1986). *The Blind Watchmaker*. W. W. Norton & Company.

Freeman, W. J. (2001). *How Brains Make Up Their Minds*. Columbia Univ Press.

Fukuma, H. (2003). A town layout and navigation. Master's thesis, The University of Tokyo.

Gescheider, G. A. (1997). *Psychophysics : The Fundamentals*. Lawrence Erlbaum Associates, Inc., 3 edition.

Gibson, J. J. (1962). Observations on active touch. *Psychological Review*, 69(6):477–491.

Gibson, J. J. (1966). The problem of temporal order in stimulation and perception. *Journal of Psychology*, 62(2):141–149.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin Company.

Grebogi, C., Ott, E., and Yorke, J. (1986). Metamorphoses of basin boundaries in nonlinear dynamical systems. *Physical Review Letters*, 56(10):1011–1014.

Hahn, J. F. (1966). Vibrotactile adaptation and recovery measured by two methods. *Journal of Experimental Psychology*, 71(5):655–658.

Hashimoto, T., Sato, T., Nakatsuka, M., and Fujimoto, M. (2008). *Recent Advances in Modelling and Simulation*, chapter Evolutionary Constructive Approach for Studying Dynamic Complex Systems, pages 111–136. I-Tech Books.

Higashiyama, A., Taniguchi, S., Miyaoka, T., and Sato, A. (2000). *Shokkaku to Itami*. Brain Shuppan.

Iizuka, H. and Ikegami, T. (2004). Adaptability and diversity in simulated turn-taking behavior. *Artificial Life*, 10:361–378.

Ikegami, T. (2007). Simulating active perception and mental imagery with embodied chaotic itinerancy. *Journal of Consciousness Studies*, 14:111–125.

Ikegami, T. and Morimoto, G. (2003). Chaotic itinerancy in coupled dynamical recognizers. *Chaos*, 13:1133–1147.

Ikegami, T. and Ogai, Y. (2010). Self-organization of subjective time and sustainable autonomy in mind time machine. In *Artificial Life XII : Proceedings of the Twelfth International Conference on the Synthesis and Simulation of Living Systems*, pages 624–625.

Ikegami, T. and Zlatev, J. (2008). From pre-representational cognition to language. In Zlatev, J., Ziemke, T., Frank, R., and Dirven, R., editors, *Body, Language and Mind*, volume 1, pages 241–283. Berlin: Mouton de Gruyter.

Jen, E. (2003). stable or robust? what's the difference? *Complexity*, 8(3):12–18.

Johansson, R. S. and Vallo, A. B. (1979). Tactile sensibility in the human hand : Relative and absolute densities of four types of mechanoreceptive units in glabrous skin. *Journal of Physiology*, 286:283–300.

Johansson, R. S. and Vallo, A. B. (1983). Tactile sensory coding in the glabrous skin of the human hand. *Trends in NeuroSciences*, 6(1):27–32.

Kaneko, K. and Tsuda, I. (2003). Chaotic itinerancy. *Chaos*, 13(3):926–936.

Kato, T. and Floreano, D. (2001). An evolutionary active-vision system. In *In Proceedings of the Congress on Evolutionary Computation*, Piscataway. CEC'2001, IEEE-Press.

Katz, D. (1925). *Der Aufbau der Tastwelt.* Johann Ambrosius Barth. Translator: Krueger, L. E., "The World of Touch", Lawrence Erlbaum Associates. Place of Publication: Hillsdale, NJ., 1989.

Konyo, M., Tadokoro, S., Takamori, T., and Oguro, K. (2000). Artificial tactile feel display using soft gel actuators. In *Proceedings of the 2000 IEEE International Conference on Robotics and Automation*, pages 3416–3421.

Konyo, M., Yoshida, A., Tadokoro, S., and Saiwaki, N. (2005). A tactile synthesis method using multiple frequency vibration for representing virtual touch. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1121–1127.

Lederman, S. J. and Klatzky, R. L. (1987). Hand movement: A window into haptic object recognition. *Cognitive Psychology*, 19:342–368.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 140:1 55.

Lund, H., Miglino, O., Pagliarini, L., Billard, A., and Ijspeert, A. (1998). Evolutionary robotics-a children's game. In *Proceedings of IEEE 5th International Conference on Evolutionary Computation*, pages 154–158.

Magee, L. E. and Kennedy, J. M. (1980). Exploring pictures tactually. *Nature*, 283:287–288.

Marocco, D. and Floreano, D. (2002). Active vision and feature selection in evolutionary behavioral systems. In Hallam, B., Floreano, D., Hallam, J., Hayes, G., and Meyer, J. A., editors, *From animals to animats 7 - The Seventh International Conference on the Simulation of Adaptive Behavior*. The MIT Press.

Miyaoka, T., Mano, T., and Fukuda, H. (1985). Threshold curves and equal-sensation contours for low frequency vibrotactile stimuli on the human glabrous skin. *Japanese Psychological Research*, 27(3):145–153.

Miyaoka, T., Mano, T., and Ohka, M. (1999). Mechanisms of fine-surface-texture discrimination in human tactile sensation. *Journal of the Acoustical Society of America*, 105:2485–2492.

Morimoto, G. and Ikegami, T. (2006). Emergence of coherent coordinated behavior in a network of homogeneous active elements. In *From Animals to Animats 9: Proceedings of the Ninth International Conference on Simulation of Adaptive Behavior*, pages 510–521.

Mountcastle, V. B., LaMotte, R. H., and Carli, G. (1972). Detection thresholds for stimuli in humans and monkeys: comparison with threshold events in mechanoreceptive afferent nerve fibers innervating the monkey hand. *Journal of Neurophysiology*, 35:122–136.

Mukaibo, Y., Shirado, H., Konyo, M., and Maeno, T. (2005). Development of a texture sensor emulating the tissue structure and perceptual mechanism of human fingers. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pages 2576–2581.

Nakajima, K., Shinkai, S., and Ikegami.T (2007). Formal model of embodiment on abstract systems: From hierarchy to heterarchy. In e Costa et al., F. A., editor, *the proceedings of ECAL 07*, pages 1110–1119, Lisbon. Springer.

Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

Nishimoto, R. and Tani, J. (2004). Learning to generate combinatorial action sequences utilizing the initial sensitivity of deterministic dynamical systems. *Neural Networks*, 17:925–933.

Nolfi, S. and Floreano, D. (2000). *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines (Intelligent Robotics and Autonomous Agents)*. MIT press.

Norman, D. A. (1981). Categorisation of action slips. *Psychological Review*, 88:1–15.

Ogai, Y. and Ikegami, T. (2007). Investigating active touch with using dinamics evolved by human selection and tactile display. In *Informatics of Dynamical Systems VI*, number Cooperative Research Reports 205. The Institute of Statistical Mathematics. in Japanese.

Ogai, Y., Shibuya, K., Ebara, H., and Ikegami, T. (2007a). Design of system generates complex system for sound softwares and installations. In *Information Processing Society of Japan Special Interest Group Technical Reports*, number 127 in 2007-MUS-73, pages 17–22. in Japanese.

Ogai, Y., Shibuya, K., Ebara, H., and Ikegami, T. (2007b). Development of sound softwares and installations using system generates complex system. In *Information Processing Society of Japan Special Interest Group Technical Reports*, number 102 in 2007-MUS-72, pages 55–60. in Japanese.

Okamoto, S., Konkyo, M., Saga, S., and Tadokoro, S. (2008). Identification of cutaneous detection thresholds against time-delay stimuli for tactile display. In *Proceedings of the 2008 IEEE International Conference on Robotics and Automation*, pages 220–226.

Osaka, N., editor (1999). *Kansei no Kotoba o Kenkyuu-suru: Giongo/Gitaigo ni Yomu Kokoro*. Shinyosha. in Japanese.

Pasquero, J. and Hayward, V. (2003). Stress: A practical tactile display system with one millimeter spatial resolution and 700 hz refresh rate. In *Proceedings of Eurohaptics 2003*.

Pfeifer, R. and Scheier, C. (1999). *Understanding Intelligence*. MIT Press.

Polhemus. Patriot 6 degree-of-freedom tracking system. `http://www.polhemus.com/?page=Motion_Patriot`. Accessed in January 27, 2009.

Port, R. F. and Van Gelder, T., editors (1995). *Mind as Motion: Explorations in the Dynamics of Cognition*. The MIT Press.

Reason, J. T. (1989). *Human error*. New York: Cambridge University Press.

Reed, E. S. (1996). *ENCOUNTERING THE WORLD:Toward an Ecological Psychology*. Oxford Universigy Press, Inc.

Reed, E. S., Palmer, C., and Schoenherr, D. (1993). On the nature and significance of microslips in everyday activities. unpulished manuscript.

Robles-De-La-Torre, G. and Hayward, V. (2001). Force can overcome object geometry in the perception of shape through active touch. *Nature*, 412(6845):445–448.

Sasaki, M. (2005). *Darwinian Method: from Motion to Affordance*. Iwanami Publisher. in Japanese.

Sato, K., Kajimoto, H., Kawakami, N., and Tachi, S. (2007). Electrotactile display for integration with kinesthetic display. In *Proceedings of 16th IEEE International Symposium on Robot & Human Interactive Communication (RO-MAN2007)*, pages 3–8.

Schwartz, A. S., Perey, A. J., and Azulay, A. (1975). Further analysis of active and passive touch in pattern discrimination. *Bulletin of the Psychonomic Society*, 6:7–9.

Suzuki, K. (2001). *Affordance and Action*, chapter Stagnation in Action transition: Microslip, pages 47–84. University of Tokyo Press. eds. Sasaki, M. and Mishima, H.

Taga, G., Takaya, R., and Konishi, Y. (1999). Analysis of general movements of infants towards understanding of developmental principle for motor control. In *Systems, Man, and Cybernetics, 1999. IEEE SMC '99 Conference Proceedings. 1999 IEEE International Conference on*, volume 5, pages 678–683.

Takagi, H. (2001). Interactive evolutionary computation: Fusion of the capabilities of ec optimization and human evaluation. In *Proceedings of the IEEE*, volume 89, pages 1275–1296.

Takahashi, H., Kajimoto, H., Kawakami, N., and Tachi, S. (2002). Electrotactile display with localized high-speed switching. In *Proceedings of the ICAT2002(12th International Conference on Artificial Reality and Tele-Existence)*, pages 10–15.

Tani, J. and Fukumura, N. (1997). Self-organizing internal representation in learning of navigation: A physical experiment by the mobile robot yamabico. *Neural Networks*, 10(1):153–159.

Thelen, E., Corbetta, D., Kamm, K., Spencer, J. P., Schneider, K., and Zernicke, R. F. (1993). The transition to reaching: Mapping intention and intrinsic dynamics. *Child Development*, 64:1058–98.

Thelen, E. and Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: Bradford Books/MIT Press.

Tsuda, I. (2001). Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems. *Behavioral and Brain Sciences*, 24:793–847.

Turvey, M. T. (1996). Dynamic touch. *American Psychologist*, 51(11):1134–1152.

Uno, R., Ogai, Y., and Ikegami, T. (2009). A cognitive linguistic analysis of onomatopoeia: Active perception in language. In *Proceedings of the 26th Annual Conference of the Japanese Cognitive Science Society*.

Varela, F. J., Thompson, E. T., and Rosch, E. (1991). *The Embodied Mind : Cognitive Science and Human Experience*. The MIT Press.

W., L. R. (2005). Dynamicity, fictivity and scanning: The imaginative basis of logic and linguistic meaning. In Pecher, D. and Zwaan, R. A., editors, *Grounding Cognition: The Role of Perception and Action in Memory, Language and Thinking*, pages 164–197. Cambridge: Cambridge University Press.

Walter, W. G. (1950). An imitation of life. *Scientific American*, 182(5):42–45.

Watanabe, T. and Fukui, S. (1995). A method for controlling tactile sensation of surface roughness using ultrasonic vibration. In *Proceedings of the 1995 IEEE International Conference on Robotics and Automation*, pages 1134–1139.