A Vibrotactile Display Based on Airborne Ultrasound Control
(空中超音波制御による振動触覚提示)

長谷川 圭介

Abstract

A technique to create an arbitrary temporal and spatial distribution of ultrasound pressure its particle and velocity in order to generate remote and non-constraint vibrotactile sensation on human body is a primary issue on this thesis. The technique takes advantage of the physical properties of ultrasound as wave propagation in terms of its high temporal and spatial resolution with good reproducibility. A nonlinear acoustic phenomenon called radiation pressure is employed in the thesis for remote non-contact vibrotactile stimuli generation on human skins. This framework has been previously proposed and called the AUTD (Airborne Ultrasound Tactile Display).

This thesis discusses the potential practical possibility of AUTD from its physical aspects. One of the goals of the study is to create a purely tangible object in the air without any appliances attached to users. For fully harnessing the property of airborne ultrasound, methods to create spatial ultrasound distributions in three-dimensional space are presented. A three-dimensional distribution of localized ultrasound is what the author thinks indispensable for generation of aerial touchable objects with volumes.

The technique described in the thesis puts its emphasis on producing vibrational sensations. It reflects the outcome of recent researches indicating that proper vibrational tactile stimuli can generate sensations of some realistic experience despite its physically incomplete reconstruction of reproduced events. A new AUTD is fabricated for generating more powerful and finer stimuli in a wider workspace compared to the previous AUTD system. The evaluation of the system is given by experiments.

AUTD used to have a problem that it generates audible noises, which might limit use of AUTD in places requiring silence. In the thesis, a generation mechanism of the noises is partially discussed and a method for its reduction is proposed and verified.

The thesis reports on two application systems of the new AUTDs, which were constructed for demonstrations in academic conferences and technological expositions, for the validation of all the achievements above. One is the 'Tactile Projector', which directly adds transparent vibrotactile sensations over the images projected on user skins. Another is the 'High-Speed Dynamic Information Environment,' which has been created in cooperation with Ishikawa-Oku Lab. in the University of Tokyo. This system can project visual images and tactile stimuli simultaneously on a hand moving freely with a special high-speed dynamic computer vision embedded in the system. Thanks to it, the projection is completed with no perceptive latency. These two systems indicate a concrete perspective of information display systems in near future.

The contribution of the technique is not only in expediting further investigations on tactile mechanism in terms of providing various tactile stimuli unique to the technique, but also widening the range of tactile applications in our daily use.

Acknowledgments

It was three years ago that I commenced researches in Shinoda Laboratory (formerly Ando-Shinoda Lab.). Not only the research scopes of the lab, but also the way Prof. Hiroyuki Shinoda interprets physical phenomenon with his own unique point of view have deeply impressed me. I greatly appreciate him for his considerable encouragements and advices based on his tremendous insights throughout the study in the thesis. Even though it is not once or twice that I have bothered him, he has supported and advised me every time I was in trouble. If there are some improvements in my way of thinking, producing new ideas and advancing researches, definitely they owe much to all what I learned from him.

I would like to express my gratitude to Dr. Yasutoshi Makino, who is a lecturer in the Lab. Since he has belonged to our lab, many inspiring discussions have been held and numerous fruitful ideas have come into being.

Research environments in the lab are what I believe that never can be provided in many other labs. The free, relaxed yet ambitious and motivated atmosphere has always filled the laboratory. The brilliant members with whom I have been engaged in researches have always helped and inspired me.

I have always respected and admired Dr. Yasuaki Monnai and Dr. Akihito Noda, both currently assistant professors in the lab, for their outstanding researches and profound knowledge. They also have been good tutors in the lab, who have supported a lot of lab students all the time.

Mr. Masahiro Fujiwara and Mr. Tatsuma Sakurai have been nice and reliable colleagues. Since they are senior to me in haptic researches, I have asked for their advise and cooperation so many times. Mr. Kazuma Yoshino is also a superb colleague whose support has lasted for two and a half years since he has joined to the lab. I believe that what I was taught by them is much more than what I taught them.

Mr. Moto Yoshioka belonged to our lab as an undergraduate student, who collaborated with me in the last year of my Ph.D. course. Several important contents in Chapter 4 and 5 in this thesis highly owe to his works.

I would like to thank my former supervisor Prof. Nobutaka Ono (currently in NII, Japan), who has established my fundamental research abilities through his lessons in my master's degree course. The research progresses in the thesis actually have much to do with the phased array signal processing technique, which I learned from Prof. Ono.

The author was financially supported by a grant from JSPS (24-9694).

Finally I would like to express my gratefulness to my family in Kasugai and Saitama, supporting and encouraging me mentally and substantially,

Thanks to people all above, I am able to continue my life as a researcher and an inventor.

Contents

Chapter 1

Introduction

1.1 Background

1.1.1 Tactile Information Processing

An indispensable goal of information processing is to expand the abilities of human modalities to their best. No matter how technologies evolve, human beings ultimately perceive their outcomes via their sense organs. Also, human beings can only affect the world around them by their actions, the movement of their muscles.

For instance, devices such as sound recorders and loudspeakers or cameras and photo printers, which store auditory and visionary information have gotten us overcoming temporal limits of modalities that we could only see and hear current occurrences. Those devices have enabled us to experience the recorded occurrences repeatedly whenever we want to. In case that stored information is mobile, we can perceive those visual and audible aspects recorded in distant places. Moreover, when electric communications are available, it is possible to sense what is happening at a different place in real time. Another kind of devices can enhance the ability of modalities. A microscope and a stethoscope are representative examples.

The author mentions the term 'information.' The most abstract concept of information has originally advocated by Shannon in 1948[1]. The theory defines information as a mathematical form. Based on this perspective, several applications such as data compression, error correction, cryptographies, signal processing and so forth, have been widely investigated and utilized. Since the information theory is highly mathematical, those applications seem something irrelevant of human sense and embodiment.

However, examples of 'information' mentioned above are related to their corresponding modalities as sights or auditory systems and are much less abstract than that defined by Shannon. In that case, information is no more irrelevant of human perceptions as it has specific names according to the correlated modalities, 'sounds' for auditory systems and 'images' for vision systems. Although it is true Shannon's information itself is represented as a sequence of numeric values, they indicate physical quantities here. They usually indicate temporal series of sound pressure for audio signals and those of coordinates in color space (e.g. well known Red-Green-Blue Coordinates).

It is no more reasonable to regard those data as irrespective of their physical and perceptual properties. For instance, audio data can be compressed into considerably small size without losing its auditory qualities by algorithms constructed through consideration of human auditory inclinations. The compression ratio here cannot be achieved by a general scheme such as Huffman Code. The same goes for video signals: e.g. compressing methods in color space preventing visual qualities of original pictures from being impaired have been of great interest among a number of researchers.

The point here is that appropriate information processing highly depends on how it is associated with human sensations. That processing is possible only when there is an intrinsic understanding over relationships between physical stimulations and human perceptions. In other words, we need to know 'what to' and 'how to' sense physical phenomenon for gaining information in order to know what is supposed to display to users as perceptive stimuli. It is also inevitable to know 'what to' and 'how to' display for generating proper sensations to users. For visions and auditory systems, studies on these relationships above have advanced to the point where daily applications of a certain practical level have become prevalent such as televisions, movies and so forth.

As described up to here, information processing methods are well established with respect to vision and auditory systems. On the contrary when it comes to the other modalities including tactile sensation, no definitive method has been founded. That fact shows that much less is known about human tactile perception. What makes studies on tactile information processing indeterminate is caused by the characteristic properties of the modality. The 'tactile sensation' is the only whole-body sensation among all human senses. Therefore the mechanism how physical stimulation is perceived as tactile sensation definitely depends on where it is. In addition, a categorization 'tactile sensation' includes a very wide range of stimuli. It physically refers to not only touch as force exerted on human bodies but also vibration induced into through body surfaces or constriction of body movements. Even a tactile cognition process limited in a small region such as a tip of an index finger, still remains a lot of mysteries unrevealed. Although it is obvious that touch sensation is aroused by mechanical interactions between the human bodies and the environment, what is physically happening at the interfaces still remains to be obscure.

These precariousness and obscurity of tactile mechanism disclosed so far have kept standard methods for tactile display and sensing from being established. As a consequence, there are few tactile applications commonly in use except a small number of examples such as notification vibrators equipped with the mobile phones. It is true that the vibrators might be designed basis on human tactile characteristics in order to minimize the energy required to create perceptive vibrotactile sensation, more detailed tactile mechanism is not taken into account. The current phone vibrators only tell the arrival of a phone call or an e-mail. More information such as who is making a phone call is still unavailable. The thesis gives a review on what has been investigated on human tactile perception so far.

1.1.2 Peculiarity of the Tactile Modality

One may think that the reason why only a small number of tactile applications are widely in use is just that visible and audible information technologies can meet our daily needs and there is no need for information display via tactile modalities any more. Here the author would like to refer to inherent traits of the tactile modality. Note that the rest of the thesis focuses on the cutaneous sensations, namely tactile sensations perceived by the skin. The term 'tactile' basically refers to what is happening on and recognized by the skin.

1. Tactile sensation is what we feel by any surfaces on our bodies.
2. We can touch and be touched by something. Tactile modality provides us with active and passive sense of touch. It is even endowed with feedback systems.
3. Usually tactile modality is more personalized than visual and auditory systems. We can share experiences of watching videos and hearing sounds concurrently with others, while it is usually impossible when it comes to tactile experiences.

These properties can be described by single words: tactile sensation is 1. whole-body, 2. interactive and 3. personal. We casually take advantage of these properties. The followings are examples with respect to them.

1. When you want to attract attention of someone who is reading a book listening to music with his head set, you can touch him by his shoulder, arm or wherever on his body.
2. You can tactually seek and locate a room lighting switch near the entrance in the dark.
3. When you are having a chat in a group and you notice that your friend sitting next to you is about to blurt out something inappropriate for the scene, you can stop him by pinching him under the table without being noticed by anyone else...if fortunate.

The first example implies that the sight and the auditory system sometimes can be occupied when engaged in some activities or tasks such as reading a book or listening to a script. On the contrary the human skin is a whole-body superficial organ, which is always ready for receiving tactile sensations even they are covered by clothes. It does not usually happen that the entire human skin is covered with hard material or occupied for some specific tasks. Thus the tactile modality is free and ready all the time. It is free and ready to be touched whether we are conscious or not. Thanks to these tactile properties, at any time we can be notified something at any time without being aware of its existence. This 'calling for attention' types of applications including mobile phone vibrators are promising. They owe much to passivity of the tactile modality, which can also be important for broadening a horizon of applications.

An intrinsic trait of tactile modalities mentioned in the second example is interactivity. This interactivity enables us to grasp an egg with force of appropriate magnitude avoiding breaking it. We human beings perform active sensing with our tactile modalities. Several properties of an object in touch with us can be detected by pushing or rubbing it with fingers and feel the response of it tactually. Thus the tactile modality can be regarded as sensors attached on moving elements of animals as actuators. It directly contributes to tactile feedbacks unnecessary for the performances of our tactile sensitivity. We can intuitively and instantaneously obtain some kind of physical information such as a small ditch on a plane. Tactile feedback on a virtual keyboard is shown to improve user performances[19], which is a good example for this. It can be stated that tactile feedback is indispensable for our accurate and agile actions in our daily lives.

The third example indicates a possibility of applications with security and confidentiality. If tactile stimuli representing some kind of information are conveyed to a person with no sound and no visible change for the other people there, that category of applications can be realized. In the design of them, it is important what kind of information is conveyed as well as how to create tactile stimuli with no interferences with the other modalities.

As stated above, tactile modality has several properties which cannot be seen in the other modalities. These properties are what we use every day for proper interaction with our external environment in a way like no other modalities. Therefore there is a manifest possibility for the advent of novel tactile applications which will be necessary for our lives in near future.

1.1.3 Difficulty in Casual Tactile Information Transmission

The author stated that more detailed understanding is essential for the prevalence of tactile applications since tactile sensation possible to be displayed is currently very limited. The

lack in the variation of generated stimuli leads to the difficulties in conveying complex information to users via tactile modalities. Yet there is another problem which hinders the spread of tactile applications. Tactile modalities need to be activated with mechanical deformation of electrical stimulation on the skin. In other words, when one intends to receive tactile stimuli via devices, the devices must be in touch with his body. They have to belong to and be worn by him. This can be bothersome once in a while. To begin with, wearing a device every time we need to receive tactile stimuli takes time. These devices oblige users to be ready for incoming tactile stimuli by being worn. In general, it is inevitable to restrict movement of users on the parts where devices mount on. The same is the case with tactile devices held by users' hands. Moreover, most of tactile devices contain complicated and bulky mechanical structures for creation of desired tactile stimuli. Eventually present tactile devices are something not tiny in size. These big, constraining devices are too inconvenient for us to wear them all through our daily lives. This situation is absolutely different from that of TVs. We do not have to wear anything for watching movies with sound. We just have to be in front of TVs and we can leave it when we want to watch it no longer. This difference is owing to the intrinsic difference between modalities. Visual and auditory modalities are sensors of incoming waves, electromagnetic waves for the former and sound wave for the latter. They take advantage of properties of a wave: it travels keeping its temporal structure through propagation, it travels much faster compared to human perception, and as the most essential one, it travels through distances. They mean devices for visual and auditory stimulation can be placed apart from users.

There is still another critical difference in tactile technologies from current visual and auditory technologies. They can cover a wide range of stimuli adequate for reconstructing scenes related to corresponding modalities with small mechanical or electrical setups. We can tell the name of a song playing from a small loudspeaker and even enjoy it. The same goes for a video. On the other hand, to the best of the author's knowledge there is no tactile display which generates tactile sensations of a variety on the same degree with a comparably simple structure. This fact directly leads to the poorer qualities of tactile sensations technically possible to reproduce than that of videos and sounds. To sum up, current tactile technology is something inconvenient yet deficient in qualities and varieties.

1.2 Goal of the Thesis

1.2.1 Two Approaches for Contributions to Tactile Applications

As mentioned above, the author stands by the side that strongly expects for the potential possibilities of tactile applications and is convinced that they will be prevalently in use and be something indispensable to our daily lives. The ultimate goal of the thesis is to solve the problems listed up above which impede the popularization of tactile applications. The author proposes two approaches for the goal:

1. Contribution to discovering fundamental mechanism of tactile information processing system.
2. Provision of tactile technologies which overcome the problem of inconvenience keeping the stimulation variety not too limited.

For the approaches, the author sets a more concrete goal of the study on creation of a new tactile display which

1. does not need to be worn by users,

2. does not constrain users' movements, and
3. can generate a wide range of tactile stimuli with their controllability of temporal and spatial characteristics.

For the first approach, there are several possible courses of handling the problem. It can be dealt with physiologically or physically. Whatever ways of problem analysis are applied, eventually the relation between exerted stimuli and their perception is of interest. In this sense it is preferable that a great deal of tactile stimuli is able to be produced. In the context of system identification, it means that we can construct a model of tactile information transmission system based on affluent input samples. Needless to say, the more input samples are available, the more precise system model can be constructed. Thus there will be no doubt that a tactile display providing diverse stimuli directly contributes to advance our understanding over the fundamentals of human tactile information processing.

Here, what kind of stimulation would be needed? The first requirement would be high temporal and spatial resolution of generated stimuli on the skin surfaces. Stimuli with spatial distributions which can drive multiple mechanoreceptors in a planer or a three-dimensional space would also be preferable. For instance, perception of rigid objects pressed on the palm has much to do with not only the responses of mechanoreceptors in regions in touch with the object, but also those in adjacent areas around these regions. This is easily understood since imaging a case above that stress concentration occurs at verges of the object. Responsive information of mechanoreceptors in a large area enclosing those in contact is necessary for the reconstruction of the tactile situation.

In addition, those stimuli should be produced with temporal feedbacks at quick responsibility. In other words, systems with high latency are not applicable. Since tactile modalities sense interaction with the outer environments, feedback latency directly deteriorates fidelity of reproduced tactile experience. A new tactile display has to possess a physical principle which realizes the tactile stimuli meeting all of the requirements stated above.

The second approach aims to removal of inconvenience derived from need for users to wear devices every time in use. As illustrated in the examples of TV, displays embedded in the environment are widely accepted and prevailed. This example describes the predominance of modalities which can handle remote physical phenomena in the context of information transmission. As the second approach, the author aims to create a remote tactile display which can be mounted on the environment providing an unspecified number of users with tactile stimuli simultaneously (Figure 1.1). Although small in number, a few approach in creating remote tactile displays have been proposed recently[9][8][15]. This fact reflects upcoming anticipation of the foundation of standard remote tactile display.

It should be noted that in recent studies on human and computer interfaces, informative systems whose function is partly extended in the environment have begun to be investigated. Harrison et al. proposed Omnitouch, an input GUI projected on walls, human bodies or any surfaces surrounding the user wearing the device composed of a three-dimensional camera and an image projector[44]. This research aims to expand regions where GUI is displayed unrestricted by the device size with this framework. Interactive use with other users can be expected. Zerroug et al. proposed the concept of 'Invoked Computing,' in which general objects around users are endowed with abilities to display vision and sound[45]. By sensing utterances of users around the objects or their body movements, the system can capture user's casual inputs as if the objects sensed them. These functions are realized by spot speakers, projectors, cameras and microphones mounted in the room. In the demonstration, bananas can be used as telephones and vacant open pizza boxes as laptop PCs. This research terminates need for users to own their own gadgets. All we need that should belong to each of us is provided by the

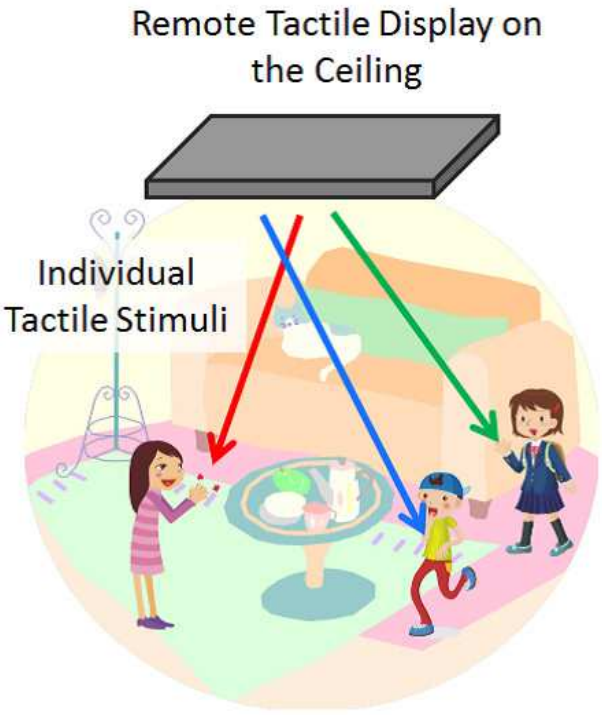

Fig. 1.1. Image of a remote tactile display embedded in the environment. It is supposed to provide individual tactile stimuli for each user simultaneously.

environment.

1.2.2 Utilization of Focused Ultrasound

It is known that focused ultrasound with sufficiently high intensity can exert static pressure on the surface of the object placed in the focal region. In order creating a tactile display which advances two approaches stated above, a focused ultrasound has been shown a promising physical principle [11][4]. Tactile displays based on this principle are called Airborne Ultrasound Tactile Display (AUTD)[4][2] (Figure 1.2). The most essential reason for employing focused ultrasound is, in short, because it is a wave. As mentioned, wave propagation has unique natures worth taking advantage of. It is prompt. Its temporal and spatial properties can be held during propagation in distance. And of course, it is remote.

Sound travels in the air at the speed of about 340 m/s. This fact implies that this principle is applicable in a tactile display system which does not allow latency violating instantaneous tactile feedbacks. Inevitable physical latency would be only several milliseconds in case the device is placed in an ordinate living room.

Sound propagation is something different from medium mass movements. It is transmission of energy through periodic movement of the medium staying the original position in a macroscopic time scale. Temporal structure of a waveform emitted at one place can conveyed to another place without being violated. This is why we can listen to words of songs traveled through long distances. The same goes for the spatial properties of sound sources. In free space, directivity of sound sources does not change through propagation. Also, phase delays at an arbitrary position in the air can be firmly determined by the dis-

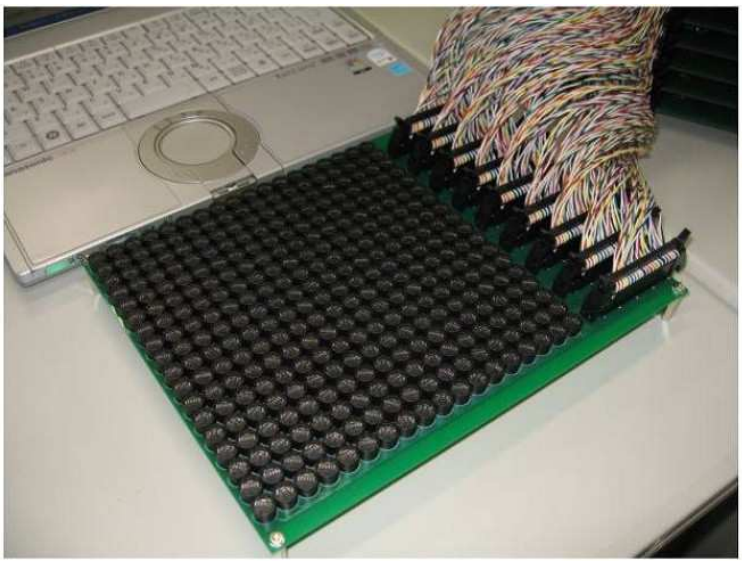

Fig. 1.2. Transducer array of AUTD developed by Hoshi et al[4].

tance from the sound source to the recorded position. This determinateness guarantees temporal and spatial reproducibility of outcoming sound field generated by waves with a determined temporal structures emitted by several sources whose position is firmly fixed. Needless to say, this reproducibility is important for applications and providing sample input stimuli for tactile studies. This cannot be achieved in principle by the air flow, which is an example of the medium movement due to unpredictability of medium stream.

Temporal structures up to several hundred kilo hertz can be transmitted without being degraded. Human tactile sensibility vanishes for vibrations faster than 1 kHz[6], where sound propagation can properly hold its temporal structure. In the thesis subtle vibro-tactile sensations are shown to be possible to generate [3].

There is one more thing worth noting. Wave field produced by multiple coherent sinusoidal point sources can shift its spatial energy peak by controlling phase shift of each element. This technique is called 'phased array technique' and has a long history. It is used in radar systems, microphone array systems, ultrasound imaging systems and so on [16][17][7]. Applying this technique to an array of ultrasound transducers, emitted ultrasound energy can localized in an arbitrary region resulting in high sound intensity causing static pressure on human bodies. This energy localization can be switched to another position by changing the set of phase shift on each transducer. The switching can be completed immediately with proper electric circuits generating driving waveform of transducers. Moreover, it has been well known that spatial distribution of sound intensity can be designed with high degree of freedom [4][41]. The thesis discusses on the possibility of creating planer or even three-dimensional tactile images in the air.

1.2.3 Contributions of the Thesis

Thus the AUTD system is promising with plenty of possibility to be a breakthrough for broadening the variety of tactile applications. The author intends to fully harness the nature of sound for widening the range of AUTD applications. The thesis reports results of several attempts to overcome technical limitations of previous AUTD system including

1. Constructing a multi-unit system of AUTD for output power enhancement and workspace enlargement.

2. Designing and displaying vibrotactile stimuli with a variety of temporal profiles.
3. Generating tactile stimulation with spatial distributions in the three dimensional region.

Physically, ultrasound focusing is done with transducer arrays playing a role as a sound lens. In an analogy to optics, it is understood that a small lens cannot produce a focused energy spot with no blur in regions far from it. In other words, the size of array definitely limits the range of the workspace where the AUTD performs effectively. A large aperture array is en essential solution to overcome this problem. Instead of creating a single large array, the author created a multi-unit array system for the sake of system setup flexibility and fabrication robustness. As a side effect, eventual increase in the number of transducers resulted in output stimuli enforcement in far field.

The effectiveness of presenting vibrotactile stimuli with high fidelity of what is recorded as sound or created through a well-constructed mathematic model is described with examples in the next chapter. Vibrotactile stimuli composed in the ways mentioned above contain multiple frequency components with their individual transient attenuation properties. Based on this fact the new AUTD system is fabricated so that it can produce vibrotactile stimuli with such a high temporal fineness. It is demonstrated that a wide variety of tactile textures can be presented by this improvement in the following chapters.

As mentioned in previous paragraphs, it is known that generating planer pattern of pressure distributions is possible with a set of transducers[4]. The author derives a method for generating pressure distributions spreading in the three dimensional space with transducers according to their locations. Although this derivation procedure is similar to other methods[41], the feasibility in the three-dimensional positions of transducers and generated sound distributions has not been verified. Moreover, the primary method are not been optimized in terms of output fidelity and its practical feasibility. The author has derived a modified method to determine output waveforms of ultrasound transducers which overcomes the problems above. As a positive side effect, mathematical regularization of the optimization problem was achieved by this modification. The method is derived with an inverse problem strategy. The author has also derived a control method of ultrasound particle velocity distribution. By controlling pressure and velocity distribution, spatial distribution of energy flow directions can be controlled. The three-dimensional energy flow distribution can create a touchable object with volume. It is possible that users touch and feel the edges, vertexes, faces or curvatures with their hands, which is unachievable with a temporal scan of a single ultrasound focus.

The author has indicated a solution for the practical problem of AUTD that it makes audible noises offensive to the ear as it generates focused ultrasound. Although this problem has been known for several years, sufficient investigations for revealing the mechanism how it occurs have not been done yet. In the thesis the author shows that one of the causes is spectral deformation due to ultrasound amplitude modulation for creating vibrotactile stimuli. Then the author demonstrates that suppression of high frequency components in the envelope waveform of modulated ultrasound reduces audible noises theoretically and experimentally.

1.3 Organization

This thesis is organized as follows. In Chapter 2, a brief review on the basic physiological tactile sensory mechanism is presented. Fundamental methods for displaying tactile stimuli are then reviewed characterized by their physical principles: mechanically-contacting, electrical and remote displays. It is also enhanced that the recent developments of vibrotactile displays and its potential usefulness. At the end of this chapter Examples of tactile

applications used or proposed are referred to. Chapter 3 describes physical principles on a sound field with its nonlinear behavior utilized in the AUTD. Sound localization strategies are described in the chapter. Sound pressure and particle velocity are the controlling targets. Mathematical optimization methods based on the inverse problem technique is derived and numerical experimental results to validate them are shown in the chapter. The actual fabricated system and its performance is presented in Chapter 4. Temporal controlling of output amplitude which is essential for creating various vibrotactile sensations is described in this chapter. Methods to vibrotactile stimuli out of recorded source are also described. Chapter 5 discusses the mechanism of audible sound generation and gives a model to describe it as a spectrum deformation by the amplitude modulation necessary for producing vibration on the skin. Based on the principle, a method to inhibit the output noise is proposed, whose effects are evaluated numerically and experimentally. In Chapter 6, two demonstration systems constructed for exhibitions are introduced. They project visible images and tactile stimuli on human body surfaces simultaneously. The research group to which the author belongs has demonstrated them several times. In the exhibitions, many participants were able to experience touchable projected objects on their palms. The author insists on the practical possibility of AUTD with these practical systems. Chapter 7 concludes the thesis and the future perspective are presented.

Chapter 2

Fundamental Principle of Tactile Information Conveyance and Reviews on Tactile Display Technology

A review for current tactile technologies is given in this chapter along with human tactile mechanism which has been investigated so far.

2.1 Nature of Human Tactile Perception

2.1.1 Categorization of Tactile Sensations

When one refers to human tactile sensations, it is usually categorized into somatic sensations. Somatic sensation is a notion contrasted with visceral sensation, which refers the sense of inner organs. Visceral sensation reflects only the internal states of human bodies. No environmental information is obtained through it.

Somatic sensation comprehends two concepts, cutaneous sensation and deep sensation. Cutaneous sensation stands for a sense on skin surfaces, in short. The human cutaneous sensory system is sensitive enough to detect a surface deformation of 10 μ m. Thanks to that sensitivity, pretty small vibrations and spatial height gaps are tactually perceptible. We are able to distinguish surface textures by touching them. Their roughness, smoothness, and even thermal conductivity can be recognized. This set of information is integrated into texture perception after processed in our brains.

A wide variety of physical phenomena perceived on body surfaces which does not constrain body movements are regarded as cutaneous sensation. Deep sensation refers to what is happening inside of the body surface. Perception of the movements of muscles, joints is classified into deep sensations. Those sensations are necessary for self-recognition of body posture, through which we can tell the stiffness and shape of an object in touch. It also contributes to detection of physical resistance against exerted force to the object, which provides essential cues for characterizing tactile properties of the object.

The thesis focuses on the former category, cutaneous sensations due to the following reasons. Required energy of physical phenomena perceived as cutaneous sensation is generally much smaller than that related to deep sensations. In the perspective of human-computer-interface applications, this property is preferable since it means that a large and bulky mechanical structure is not always necessary for the fabrication of cutaneous-stimuli-presenting devices. In addition, human skin is the outermost organ and literally is a margin in touch with inside of human beings and the outer world, which is a natural 'interface.' Therefore well-organized stimuli with realities on human skin can provide us

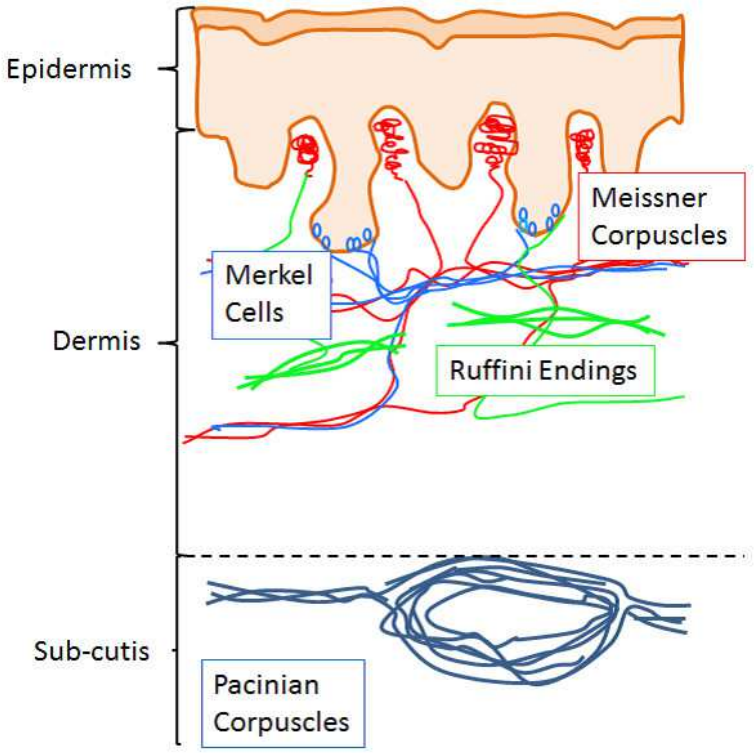

Fig. 2.1. Structure of the cutaneous mechanoreceptor system. Reconstructed from [12]

with affluent information as a highly intuitive form.

Of course it does not means that deep sensation is not an important study subject. The wholly understanding over what is considered 'haptics' inevitably includes that of deep sensation. Moreover, a great deal of tactile events is perceived via both of cutaneous and deep sensation. Here the author would like to emphasize those investigations on deep sensation still remains as one of indispensable and vast field of haptic researches.

2.1.2 Cutaneous Sensations

There has been a pile of studies on cutaneous sensations. Early researches identified tactile mechanoreceptors in dermis and subcutis, which is covered by epidermis. The thickness of epidermis is approximately 0.7 mm and that of dermis is around 1∼3mm. Thus these mechanoreceptors are placed at a depth of several millimeters. They react to mechanical stimuli around themselves which are converted into centripetal electric signal pulses.

Beside mechanoreceptors, there are several sensory receptors correlated to cutaneous sensations. Thermoreceptors and nociceptors contribute to perception of temperature and pains, respectively. It has been found that there are polymodal receptors which receive multiple kinds of information. In the thesis those sensory receptors are not described any further.

Anatomically, four classes of mechanoreceptors have been found. They are Meissner corpuscles, Merkel cells, Ruffini endings and Pacinian corpuscles. They are classified into Slowly Adapting (SA) mechanoreceptors or Fast Adapting (FA) mechanoreceptors (Figure 2.1). SA receptors emit intermittent nerve impulses for minutes since the initial contact on the skin, while FA receptors stop firing impulses soon after the contact[12]. Another

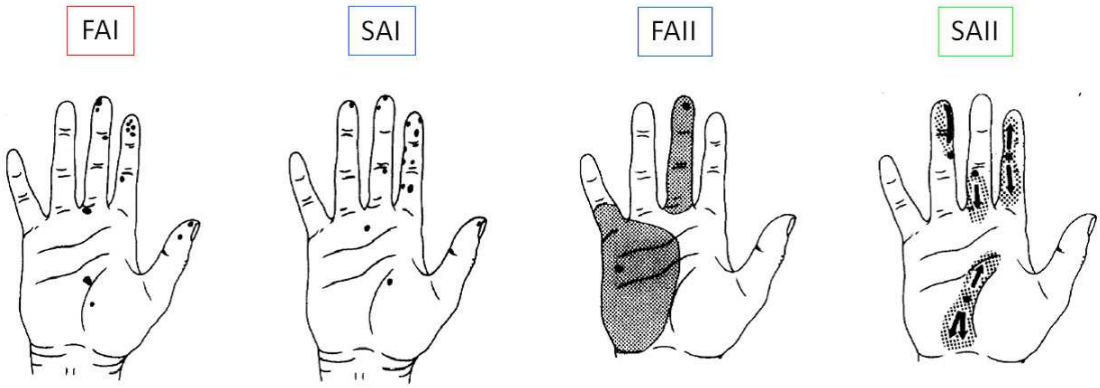

Fig. 2.2. Receptive fields of four mechanoreceptors. Dashed regions indicate the ranges. Rearranged from [21].

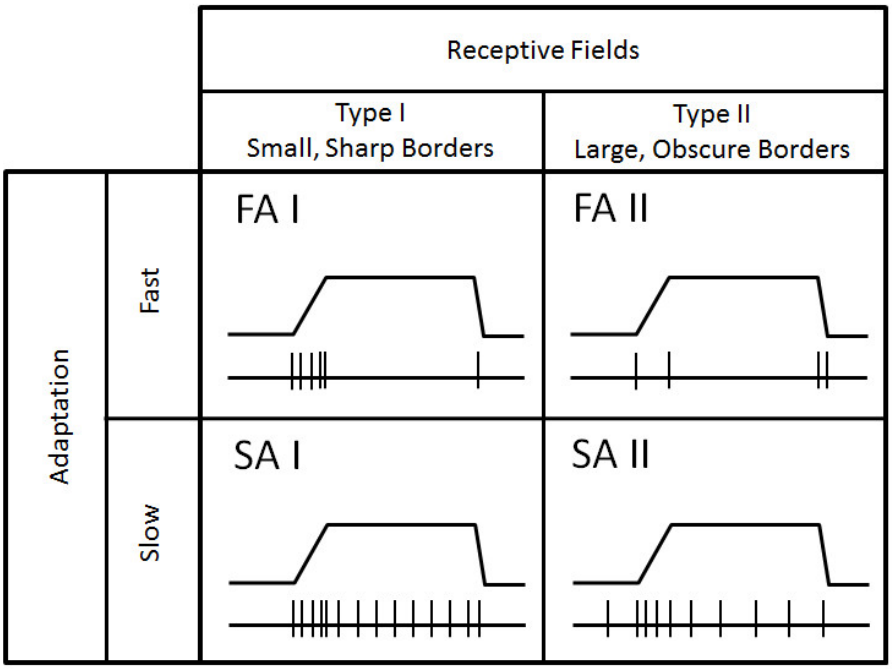

Fig. 2.3. Temporal adaptation inclinations of four mechanoreceptors. Upper graphs indicate the skin displacement and the lower indicate the timings of nerve impulse firing. Reconstructed from [12].

aspect of classifying them is the spatial ranges in which each of them responds to exerted stimuli[21]. The type I receptors covers small regions around them whose borders are explicit. The type II receptors react to stimuli in much larger regions with opaque borders (Figure2.2). These two standards, temporal reactivity and sizes of receptive fields characterize the four mechanoreceptors above. They are occasionally indicated as following; the Merkel cells are indicated by (SAI), Ruffini endings by (SAII), Meissner corpuscles by (FAI) and Pacinian corpuscles by (FAII). Their temporal responses are schematically depicted in Figure 2.3. From their adaptation characteristics, their functional trait is un-

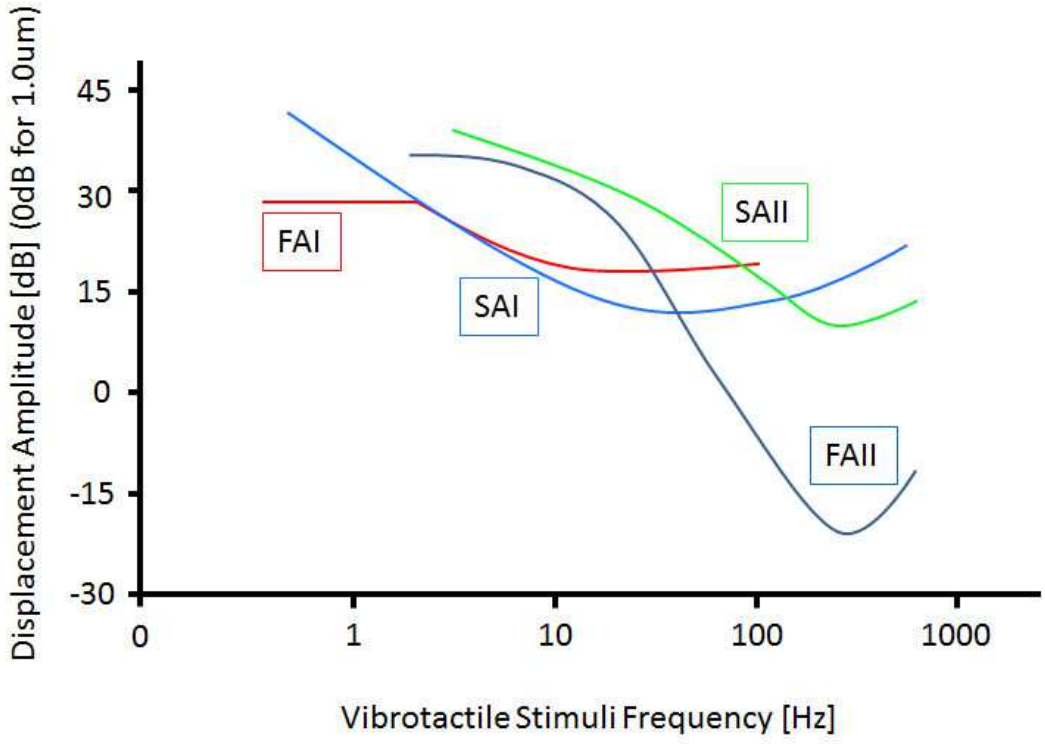

Fig. 2.4. Displacement thresholds of four mechanoreceptors with respect to vibration frequencies. Reconstructed from [20]

derstood as following. SA mechanoreceptors have slow adaptation speed i.e. nerve their impulse firing lasts for minutes. Hence they response to deformation with very slow time scale, almost static pressure imposed on the skin. Difference between SAI and SAII partially remains unrevealed since literatures handling with SAII properties are so small in number. From their differences in receptive field range, SAI receptors are said to capture skin curvature or spatial pattern of exerted pressure on the skin. It is believed to some extent that pinching or pulling of the skin is partly dealt with SAII receptors. Functional characteristics of FAI and FAII are more clearly understood. The most outstanding difference is that FAI responses to velocity of deformation while FAII does to deformation acceleration. This difference is owing to their mechanical structures.

Their difference in temporal adaptation inclination can also be interpreted as their inherent frequency characteristics in sensitivity of vibrotactile sensations [20]. Figure 2.4 schematically depicts the threshold vibrational displacement of four mechanoreceptors within the frequency range of 0∼1000Hz. It is obviously seen that FAII receptors indicate the most sensitiveness among all receptors at around 250Hz. Its displacement threshold is nearly 0.1μm at 250Hz. FAI has its sensitiveness peak of 6 μm threshold at 40Hz. SAI has peak frequency around 50Hz. Displacement thresholds SAI is 10μm. There is still much obscure about the behavior of SAII. Its threshold is believed to be around 3μm and frequency response is similar to that of SAI. However, we can see that contribution of SAII receptors to sinusoidal vibrotactile sensitivity cannot be seen in Figure 2.4 because at all frequencies SAII is not the most sensitive.

Based on the frequency selectivities among mechanoreceptors, a method for synthesizing vibrotactile sensations the way colored images were composed has been originally

proposed by Shinoda et al.[13]. It is widely known that a full-colored image can be created with at least three overlapping luminous images of different colors such as red, green and blue. This is owing to the physiological knowledge that human retinas possess arrays of three types of light receptors with their individual frequency responses. The same mechanism is seen in human mechanoreceptive systems as described above. A sinusoidal stimulus which selectively evokes a specific class of mechanoreceptors keeping the others excited at negligible degrees can be the basis stimulus to that mechanoreceptive channel. In an analogy to the image reconstruction, it is expected that tuning temporal envelopes of multiple basis stimuli creates a great variety of vibrotactile sensations. Shinoda et al. take advantage of the depth difference of mechanoreceptor locations for selectively stimulating a specific class of mechanoreceptors with mechanical vibration. Induced vibration attenuates as it goes through depth. By taking account of this filtering effect of the skin, induction depth of vibration can be controlled. This strategy is widely used in designing vibrotactile sensations.

2.2 Tactile Display Technologies

2.2.1 Overview

In this section the author gives a review on tactile display technologies investigated so far. Some of the investigation here includes state-of-the-art achievements.

There have been numerous tactile displays proposed. In general they are designed to generate specific aspects of tactile stimuli. Hence they can be categorized by provided tactile sensation such as superficial textures, fineness, vibrations, and so forth. There is another way of classification of tactile displays according to how they create tactile sensations. In the thesis, tactile displays based on mechanical contacts with human skin are reviewed at first in the next subsection, where the author emphasizes on the vibrotactile displays in the latter part. Brief reviews on electric tactile display technologies and remote tactile displays are given in the following subsections.

2.2.2 Mechanical Tactile Displays

Most of current tactile display produces tactile stimuli by creating mechanical deformation on the skin in touch with them. One of the most fundamental types of tactile displays is pin-array tactile displays. They are composed of movable pins disposed in an array. Two-dimensional arrays are common among pin-array tactile displays. The pins move vertically to the array plane in many devices[22][23]. In some devices the pins move laterally and it is reported that this lateral vibration generates the feeling that pins push the skin surface vertically[26].

There is another remarkable research reporting the effect of creating tactile stimuli with suction pressure[46][47]. In the investigation above, it was experimentally and theoretically verified that we tactually feel that a region on the palm where suction pressure is imposed, in short being pulled by vacuum, is pressed with pressure. This perceptive phenomena is analyzed by Makino et al. and concluded that the cutaneous strain energy distribution becomes similar in case the skin surface is pushed by a solid object and in case sucked by vacuum, which gives similar sensations. In recent tactile displays this mechanism is applied to for presenting griping force at the fingertip[48]. MEMS techniques are also introduced in some tactile displays [49]. They differ in the maximum force possible to apply, spatial resolution of stimuli or temporal responses.

Those displays are supposed to reconstruct deformation on the skin as users touch

objects. There are yet many researches focusing on the presentation of other tactile aspects. Nara et al. fabricated tactile displays using surface acoustic wave (SAW)[24]. The display consists of a substrate on which the SAW travels and a slider on it composed of a thin plate covering steel balls. User putting their finger on the slider can feel the fineness under the slider varies by controlling the friction magnitude with the SAW. This is an example of displaying variable granular textures. It is created by controlling not superficial deformation at individual point of the skin but the whole physical phenomenon between the finger and the device. Minamizawa et al. created GravityGrabber, which adds tactual feeling of variable virtual weight on the fingertip gripping a generic object[25]. The device is composed of a sliding band in contact with a finger cushion which is connected to a pair of motors. The movement of motor is translated to the shear force applied on the finger tip. When user grabs a generic object through this band of the device, the additional shear force is perceived as the weight variation. This device is also able to display time variant virtual inertial mass with proper feedbacks. The device is free from large mechanical structures.

A full-body tactile display is also an object of researchers. Lindeman et al. has proposed a whole body tactile display system which can be used with body movements of users. Optical motion capture system is combined with the system so that users can freely move in virtual worlds reflecting the feedback of their body movements. The authors mention that luck in qualities of displayed tactile sensation remains as one of future improvements since it is created with a set of vibrators with on-and-off control. Yet this device is not actually 'whole-body' because there are surfaces not covered by the device. It is pretty tough challenge creating a whole body tactile display which allows a completely free movement of users. Normal wiring in devices does not permit highly stretching and twisting deformation caused by user movement. Shinoda et al. has demonstrated that wireless communication with electromagnetic waves bounded in layers made of elastic and stretchable materials and that it can be applied to soft robot skin [50]. The key technique is called 'two-dimensional communication[51][52]', which could be a promising framework for creating a whole-body tactile display, although still it is a challenging work.

2.2.3 Recent Vibrotactile Displays

In the following example the author would like to emphasize that a cutaneous vibrotactile stimuli convey affluent information for the tactile perception of many physical aspects.

Hayward et al. demonstrated that a device composed of a rod with a vibrator mounted on it can present a feeling of a rolling stone moving inside the device according to its posture to users holding it[31]. The presented vibration is well-organized and varies as the rod posture, which is sensed by the accelerometer installed on the device for detecting the ambient gravity force, is changed. Users can perceive the virtual rolling stone moving back and forth by inclining the device. The point here is there is no internal change of mass distribution in the device. Although it is just vibrations what a user is feeling, he can clearly perceive and picture the virtual movement of the stone. This fact implies that there is no need for the reproduction of the entire physical phenomena which are supposed to be happening at the time of a tactile experience in order to reconstruct it. This principle helps tactile applications spread in use because in many cases it will be possible to reduce the mechanical complicities of devices if they are supposed to produce properly simplified mechanical movements. Here it is important investigating what is the most intrinsic physical events arousing perceptions of a specific tactile experience.

The 'TECHTILE Toolkit' is a novel approach to widen and investigate the extent that the tactile experiences which vibrotactile sensation essentially contributes to their intuitive perception[32]. Its system setup is quite simple. A set of microphone and a vibrational

actuator with a power amplifier is what it comprises. The recording microphone can be attached to various generic objects such as a paper cup, a tennis racket, a plastic bag and so forth. When they are attached to objects, they record the vibration of the objects as sounds. The recorded vibration is then stored, or transferred to the actuators with properly amplified power. The actuators also are attached to objects, usually the same ones with which corresponding microphones are in touch. Then the vibration is transmitted from one object to another, which is in some cases effective for representing the physical experience with tactile modalities. Rolling balls or poured water in a paper cup, potato chips moving in the shaken plastic bag are well perceived by the system. There are advanced researches using multiple microphones and actuators as well[42]. It should be mentioned that this system can reproduce some physical phenomenon with high fidelity but there are other cases it does not work well. For example, it is easily imagined that gradual mass increase with small vibrations such as water poured gently into a cup is not well reconstructed with the system. Yet it is worth noting that the system revealed that in some cases spatial distribution of the vibration does not play any roles for tactile perception of the phenomenon. However since it is just a 'toolkit', the theoretical analysis of tactile perception with the system is not so far advanced.

A challenge in constructing a stimulus model in vibrotactile sensation reconstruction in recent researches is seen. Kuchenbecker et al. proposed a method to create a numerical model to reconstruct textures of material surfaces by vibrotactile sensations[33]. A pen-shaped device is utilized for recording the vibration of the surface bump of a rigid planer material. The sampling of materials is done by moving this device on the material for seconds as vibrations. Stochastic modeling techniques are employed for creating the vibrotactile stimulus model from sampled vibrations. In the reproduction phase, reconstructed vibration is presented via a pen-type device. The current version has two primary limitations: the material has to be isotropic and the contact with the material has to be made via an apparatus with a single point. Nevertheless it is expected that from this modeling method vibrational information necessary for texture perception is revealed.

2.2.4 Electrical Tactile Displays

Up to here examples in mechanical tactile displays are shown. From this subsection tactile displays based on other principles are introduced. Since tactile sensation is aroused by the nerve impulses arriving at the brain, it is natural to think of produce it by proper temporal or spatial patterns of electric pulsed injected in the afferent nerve system. An approach to excite nerves related to tactile perception has a long history[27][28], though in early researches it was not a research scope to generate a tactile stimuli with natural texture at high qualities. Kajimoto et al. proposed a physical model about generation of electrical tactile sensations in relation to the principle of tactile primary colors mentioned above[13] and verified its validity with experiments[29]. Improvement in displayed tactile sensation quality is also demonstrated to be possible by taking into account the electrical impedance change on the skin monitored[30]. This electro-tactile display does not require electrodes injected under the skin, but superficial electrodes in touch with skin surfaces. This principle is introduced in related applications[14][53]. The merit of using electric stimuli on skin surfaces is that compact, lite and low power-consuming devices can be fabricated. There is still a remaining challenge in terminating the painful sensation in tolerable for comfortable use.

2.2.5 Non-contact Tactile Displays

The author must state that there are a small numbers of remote tactile displays proposed. The current remote tactile displays including AUTD use the air for medium to convey tactile stimuli to users since it is the only material always in touch with human body surface regardless of their postures. Remote tactile displays beside AUTD make use of the air flows, which is the most essential difference from AUTD. Tactile displays using the air jet has been proposed by several researchers[9][8]. They can be applied to the situations where a direct touch to the device is not preferable by some reasons such as hygiene requirements. However, there are inevitable problems against the convenient use in applications owing to the physical nature of the air. First, the air jet displays might be remote but not really. To be precise, they can be said to be 'non-contact' instead of 'remote'. Most of the air coming put from the jet nozzle can not reach a desired region far from the nozzle. This is simply because the air spreads in every direction as soon as it is exhaled. However high the stream velocity at the nozzle might be, at distant region it is highly attenuated. Another critical problem is that the temporal and spatial resolution too degraded in far field. Owing to the spreading behavior of the air, it is quite difficult to display localized sensation patterns on human skin. The temporal profile is also hard to control. In addition, when the stream velocity becomes high, the behavior of the flow becomes unpredictable. This is critical for applications required high reproducibility. As mentioned, the air flow is usually very difficult to control with high resolutions.

Another research example takes advantages of other properties of the air. The device proposed by Sodhi et al. uses an air vortex, which travels with higher directivity compared to the air jets[15]. By constructing a vortex, condensed energy of emitted air can travel much longer distances. This trial seems to overcome the problems of the previous examples. Yet still, remaining problems exist. First, the velocity of a vortex with the system is about 7.2 m/s causing latency of hundreds of milliseconds to users one meter apart from the device, which is easily detectable for us. The size of vertex is about 8.5 cm in diameter. Physically, the smaller the vertex size becomes, the nearer it can reach. This fact indicates that the system cannot generate tactile sensations with high spatial distribution in far field. Moreover, behaviors of a vortex are still unpredictable. Nevertheless, it should be remarked that the device is small in size and have simple mechanical structures including loud speakers. Applications for mobile use with the system are promising.

2.3 Tactile Applications

In this section examples of tactile application are shown along with how tactile modalities are and are supposed to be used.

2.3.1 Sensory Substitution

The term sensory substitution refers to the activity of a modality covering the function of other ones which is not originally allocated to the one, often due to the disabilities of original sensory systems. Since the tactile modality is a full-body sensor, there is plenty of room in it for sensory substitution. The Optacon (OPtical to TActile CONverter) is one of the most famous sensory substitution device[34][35]. It can convert optical brightness of printed letters into tactual bumps. In short the device intends to enable us, especially for the visually handicapped, to read sentences with our touches. Visual modalities are expected to be substituted by tactile modalities. It is reported that well-trained users

can read 40 words a minute with the Optacon and some can read about 100 words a minute[54], which is half of the speed the well-trained braille reading[55][56]. Though the training aside reading brailles is necessary, it is useful in perception of written shapes. Devices translating sound to tactile sensations are also undergoing research subjects[57].

2.3.2 Job Training

Tactile training is essential for some occupations. People engaged in tasks requiring dexterous operation must learn how to associate possible situation with received tactile feedbacks. There are several surgery simulators endowed with tactile feedback displays in professional use[36][37]. Those simulators give tactile feedback to mimicry surgical instruments in the realistic situation being simulated. Although the chances to train surgical skills with real patient are limited, use of those simulators provides limitless training opportunities with no risks.

2.3.3 Tactile Experience in Amusement Attractions

In many amusement attractions, use of modalities beside sights and auditory systems has been accepted. 'Micro Adventure!' in the Tokyo Disneyland is an old example (now it does not exist). Recently, those attractions are referred to '4D' attractions. Apart from the general definition of dimension, '4D' here refers to four modalities, sight, sound, scent, and touch. In general they are embedded in the theatre so they are often called '4D-films'. The touch plays an important role in those films. In Vancouver Aquarium, a film with marine creature is showing [43]. Customers sitting back on the seats suddenly feel being poked by a fish as the movie shows the poking fish or electric vibration of an electric eel. Those stimuli are generated with mechanical structure embedded in the seat. For this kind of applications, it has to be guaranteed that postures of users are almost determined during the operation of the attraction.

In many gaming controllers vibration has been used for enhancement of sense of presence for a long time. In some platform they react to the body movement of users for providing a kind of tactile feedbacks[58]. Though poor in tactile quality, they contribute to the excitement of players.

2.3.4 Tele-existence

Tele-existence is a concept that a robot (often called 'Slave') in a different place moves and acts as the user (called 'Master') does providing the whole sensory stimuli received by the Slave robot where it is placed[59]. The ultimate goal of the tele-existence system is to create a remote alter presence of a user so that the user can feel the world at a different place and affect to it as if he were there. Researches of tele-existence have a long history and are still ongoing. Current tele-existence system 'Telesar V' provides tactile feedbacks classified into cutaneous and deep sensations[38]. Master users can touch and feel objects via the slave robot.

Touch sensations in tele-existence systems is indispensable in many situations. For examples, in case of the operation of dangerous and difficult tasks through the system like repairing a nuclear plant, the dexterousness of the user directly depends on the tactile feedback quality provided. Because there are a kind of physical information which is comprehended most intuitively and quickly by touch. This concept of tele-existence is something enshrined in the center of virtual reality, where Tactile technology also takes one of the most essential places.

Chapter 3

Localizing Ultrasound with High Intensity

This chapter describes fundamental properties of localized ultrasound. They are indispensable for proper modeling of physical phenomena which are harnessed for a tactile display.

3.1 Nonlinear Behavior of Sound Field

3.1.1 Linear Cases

The Airborne Ultrasound Tactile Display employs nonlinear behaviors of a sound field. Usually, a sound field is governed by a linear wave equation. Owing to the linearity, time-harmonic handling of wave equation is guaranteed. This linearity is based on an assumption that relative change in pressure is sufficiently small. Sound pressure p[Pa] is defined as a perturbation of pressure from the static air pressure P_o[Pa] where there is no sound. When the sound pressure under the existence of sound whose amplitude is p becomes P[Pa], their relation yields:

$$P = P_0 + p \tag{3.1}$$

Where this equation holds, the assumption above, $|p| \ll P_0$, leads to

$$\frac{|p|}{P_0} \ll 1. \tag{3.2}$$

Here the term $\frac{|p|}{P_0}$ is the relative change in pressure. A system of differential equations describing the behavior of ideal gas with no viscosity is given as following [60]:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0 \tag{3.3}$$

$$\rho \left\{ \frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} \right\} = -\nabla P_0 \tag{3.4}$$

$$P = P_0 \left(\frac{\rho}{\rho_0} \right)^{\gamma}, \tag{3.5}$$

where ρ_0[kg/m^3] denotes the medium density where there is no sound, $\boldsymbol{v}$[m/s] denotes the particle velocity and γ[-] denotes the heat capacity ratio of the medium. Note that all of those parameters $\rho, \boldsymbol{v}, P$ and p are represented as functions of time and location as $p(\boldsymbol{r}, t)$

by letting $\boldsymbol{r}$ be the location in the medium and t be the time. (3.3), (3.4) and (3.5) are called the continuity equation, Euler' s equation of motion and the adiabatic equation. It should be noted that this system of equation holds in non-linear acoustic fields.

In many cases, sound amplitudes $p(\boldsymbol{r}, t)$ and particle velocity $\boldsymbol{v}(\boldsymbol{r}, t)$ is handled with another physical parameter $\phi(\boldsymbol{r}, t)[\mathrm{m}^2/\mathrm{s}]$ called velocity potential, whose definition is given by

$$\boldsymbol{v}(\boldsymbol{r}, t) = -\nabla\phi(\boldsymbol{r}, t). \tag{3.6}$$

From the assumption (3.2), (3.3),(3.4) and (3.5) are rewritten as

$$\frac{\partial\rho'}{\partial t} + \rho_0\nabla\cdot\boldsymbol{v} = 0 \tag{3.7}$$

$$\rho_0\frac{\partial\boldsymbol{v}}{\partial t} = -\nabla p \tag{3.8}$$

$$p = c^2\rho' \tag{3.9}$$

by terminating infinitesimal terms with the second orders. Here $\rho'[\mathrm{kg}/\mathrm{m}^3]$ is a density variation from ρ_0 holding a relation $\rho = \rho_0 + \rho'$. $c[\mathrm{m}/\mathrm{s}]$ denotes the sound velocity. It is also confirmed that the equation $\nabla\times\boldsymbol{v} = 0$ holds and thus a linear sound field is an irrotational field. Hence the sound amplitude and the density variation can be integrated as a function of ϕ:

$$p = p_0\frac{\partial\phi}{\partial t} \tag{3.10}$$

$$\rho' = \frac{\rho_0}{c^2}\frac{\partial\phi}{\partial t}. \tag{3.11}$$

Eventually the velocity potential $\phi(\boldsymbol{r}, t)$ is shown to hold the following equation

$$\nabla^2\phi - \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = 0. \tag{3.12}$$

This equation is called as the wave equation since its general solution is given by the form of propagating wave $\phi = A(\boldsymbol{r})e^{j\omega t}$. The most remarkable property of linear sound field is that it holds superposition principle. In other words, the linear combinations of multiple waves with their individual frequencies and phase delay hold the wave equation. This property makes the Fourier analysis feasible.

The author introduces physical description of linear wave fields in spite of that the nonlinear behavior of high pressure sound field is an intrinsic principle of AUTD. This is because almost all regions in a sound field created by AUTD can be described by the linear sound field theory due to its low sound amplitudes. A region where condensed energy exists and nonlinear sound behavior is seen is limited. It has been shown that the phased array technique with linear wave field accounts for the resulting sound amplitude field [4][3].

3.1.2 Acoustic Radiation Pressure

With high amplitude, nonlinear behavior of sound field is seen such as waveform distortion (Figure 3.1), acoustic streaming and acoustic radiation pressure, used in AUTD. This is understood as static pressure appearing as the system of equations (3.3) $\sim$ (3.5) is handled with second-order approximation.

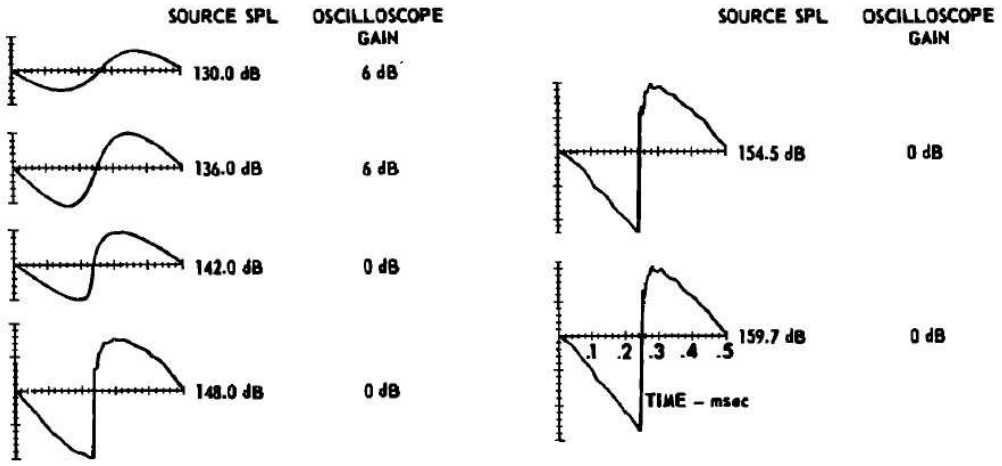

Fig. 3.1. Waveform distortion at high amplitude. Reconstructed from [66].

Let us begin with series defiled as the following:

$$P = P_0 + p_1' M + p_2' M^2 \tag{3.13}$$

$$\boldsymbol{v} = \boldsymbol{v}_1' M + \boldsymbol{v}_2' M^2, \tag{3.14}$$

where M[-] is the Mach number of the medium. The coefficient p_1'[Pa], p_2'[Pa], $\boldsymbol{v}_1'$[m/s], $\boldsymbol{v}_2'$[m/s] determined by the values of P and $\boldsymbol{v}$ i comparison with M. If P and $\boldsymbol{v}$ is smaller enough, the second order coefficients $p_2', \boldsymbol{v}_2'$ become zero. By substituting these series to (3.3) $\sim$ (3.5), it is derived that

$$< p_1 > = 0, \tag{3.15}$$

$$< p_2 > = \frac{\rho_0 |\boldsymbol{v}_1|^2}{2} - \frac{p_1^2}{2\rho_0 c^2}, \tag{3.16}$$

where $p_1 = p_1' M$[Pa], $\boldsymbol{v}_1 = \boldsymbol{v}_1' M$[m/s][62][5] and $< \cdot >$ denotes the time average operator. The value of $< p_2 >$ becomes zero when there is no object blocking the traveling wave propagations. However in case that dissipative objects placed in the sound field, energy density gap between the outside and inside of the objects. This density gap causes pressure called 'acoustic radiation pressure.' The magnitude of radiation pressure P_R[Pa] can be estimated by the following equation:

$$P_R = \alpha \frac{< p^2 >}{\rho_0 c^2}, \tag{3.17}$$

where α[-] is determined by $\alpha = 1 + R^2$ with R[-] denoting the reflection coefficient on the object. The value of α is from 1 to 2. In case wave inducing a rigid object from the air, most of the energy is reflected on the object surface. As a result α becomes nearly 2. This situation is different from the radiation pressure tactile display fabricated in the previous studies using water as the propagating medium[64].

The acoustic pressure defied above contains the time average of the square of sound amplitude, which is in proportion to sound intensity. Temporal change in radiation pressure must be argued in a longer time scale than the temporal cycle of the sound pressure. AUTD exerts static or time-variant force on human skin surfaces by utilizing acoustic radiation pressure. Note that the pressure exerts on the surface with its normal direction.

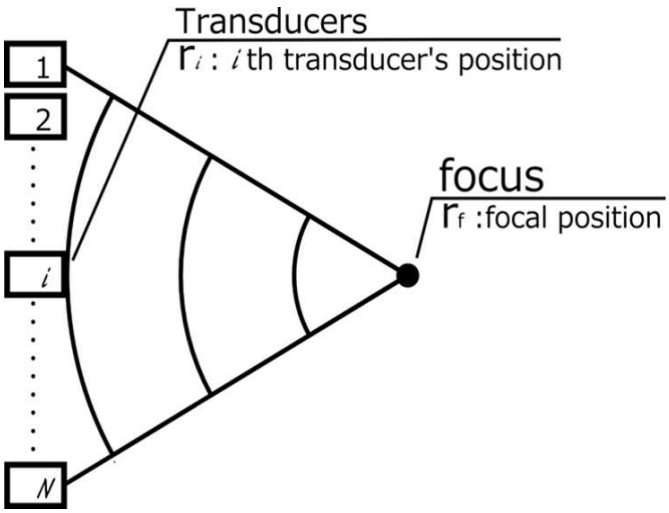

Fig. 3.2. Schematic image of a phased array focusing.

No shear force has been found to be generated by this principle. This nonlinear acoustic phenomenon arises in the sound field with the intensity of 120dB∼[60].

3.2 Generating Amplitude Field with Acoustic Phased Array

In order to generate acoustic radiation pressure intense enough to be perceived tactually, a high amplitude sound field must be produced. It is also important to generate localized sound intensity for producing tactile sensations with spatial resolutions. In order to meet the two requirements above, the phase array technique is introduced in the AUTD system[2][4]. The phased array technique is an energy localization method with a set of coherent wave sources whose individual phase shifts are tunable[16].

3.2.1 The Case with a Single Spot

Let us begin with the case where a single focal spot of sound intensity should be generated. The author assumes that the positions of ultrasound transducers are given. This assumption accounts for the situation that the devices are embedded firmly in the workspace. The other assumption is that they emit coherent sinusoidal sound of a common frequency $f = \omega/2\pi$[Hz]. Let $\boldsymbol{r}_i$[m] be the 3-dimensional position of the i-th transducer among N transducers equipped to the AUTD and $\boldsymbol{r}_f$[m] be the desirable 3-dimensional (including the depth of the focus) focal position. This problem definition is schematically depicted in (Figure 3.2).

Let θ_i be the phase shift added to the i-th transducer. Here, the output sound from i-th transducer $p_i(t)$ is given by

$$p_i(t) = p_t e^{j(\omega t + \theta_i)}, \tag{3.18}$$

where p_t[Pa] denotes the common output amplitude of transducers. In order to maximize the sound amplitude at the focal point $\boldsymbol{r}_f$, phase shift on each transducer should be determined so that it compensates the phase delay through distance due to the finite sound velocity:

$$\theta_i = k|\boldsymbol{r}_i - \boldsymbol{r}_f|. \tag{3.19}$$

Here $k = \omega/c$[1/m] denotes the wavenumber.

Once the transducer phase shifts are determined, the resultant wave field $p(\boldsymbol{r})$ is given by the superposition of their output sound:

$$p(\boldsymbol{r}, t) = \sum_{i=1}^{N} a(\boldsymbol{r}_i, \boldsymbol{r}) e^{-jk|\boldsymbol{r}_i - \boldsymbol{r}|} p_i(t) \tag{3.20}$$

Here $a(\boldsymbol{r}_i, \boldsymbol{r})$ is an attenuate coefficient determined by spatial conditions which can usually be decomposed into several factors. The most common inevitable attenuation is due to the inverse-square law of intensity propagation. According to this physical law, attenuation factor should include

$$a_1(\boldsymbol{r}_i, \boldsymbol{r}) = \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}|}. \tag{3.21}$$

Next, frequency characteristic loss of the air is described as

$$a_2(\boldsymbol{r}_i, \boldsymbol{r}) = e^{-\beta|\boldsymbol{r}_i - \boldsymbol{r}|}. \tag{3.22}$$

From [65], the attenuation by a_2 is estimated as 100dB/200m for 40kHz ultrasound transducers. The factor β is proportional to the square of the sound frequency. The other factor would be the directivity of transducers. It should be expressed as a function of angles determined by the direction from the transducers. The other factor $e^{-jk|\boldsymbol{r}_i - \boldsymbol{r}|}$ denotes the physical phase delay though propagation. Since the temporal factor is extrinsic in the situation, by omitting it (3.20) can be rewritten as

$$p(\boldsymbol{r}) = \sum_{i=1}^{N} a(\boldsymbol{r}_i, \boldsymbol{r}) e^{-jk|\boldsymbol{r}_i - \boldsymbol{r}|} p_i \tag{3.23}$$

with $p_i := p_t e^{jk\theta_i}$.

The equation (3.23) is not intuitive for comprehending how the resulting sound field would be. Hoshi et al. gives a approximation of amplitude distribution in case that a planer square array in which transducers are allocated in lattice whose interval d_a is constant[4]. In the plane parallel to the array at the depth of focal position $\boldsymbol{r}_f = [x_f \ y_f \ z_f]^T$, the amplitude distribution is given by

$$p(\boldsymbol{r}) = p_r N^2 \frac{\operatorname{sinc}\left(\frac{kd_aN(x_f - x)}{2z_f}\right) \operatorname{sinc}\left(\frac{kd_aN(y_f - y)}{2z_f}\right)}{\operatorname{sinc}\left(\frac{kd_a(x_f - x)}{2z_f}\right) \operatorname{sinc}\left(\frac{kd_a(y_f - y)}{2z_f}\right)} \exp\left(k\frac{x_f^2 + y_f^2 - x^2 - y^2}{2z_f}\right). \tag{3.24}$$

Here the representative amplitude p_r is determined based on an assumption that the attenuation coefficient $a(\boldsymbol{r}_i, \boldsymbol{r})$ can simply be expressed as $a = a_0(z_f)$. The origin of coordinates is set to the center of the array. (3.24) is much more intuitive to understand. The amplitude density varies according to the product of two orthogonal sinc function with two axis. With this expression we can estimate the spot size as the distance between the two central zero cross $\frac{4\pi z_f}{kNd_a}$. It is also implied that this is a periodic function with respect to x, y, which indicates the existence of grating robes. A numerical simulative example of amplitude distribution near the focal point is shown in (Figure 3.3) [4]. It should be noted that the focusing performance is degraded in the region far from the radiation plane. This phenomenon is understood in an analogy to optical focusing, in which the size of lens directly confines the range where proper focusing can be produced. As a focusing performance of a lens highly depends on its aperture size, the array size of

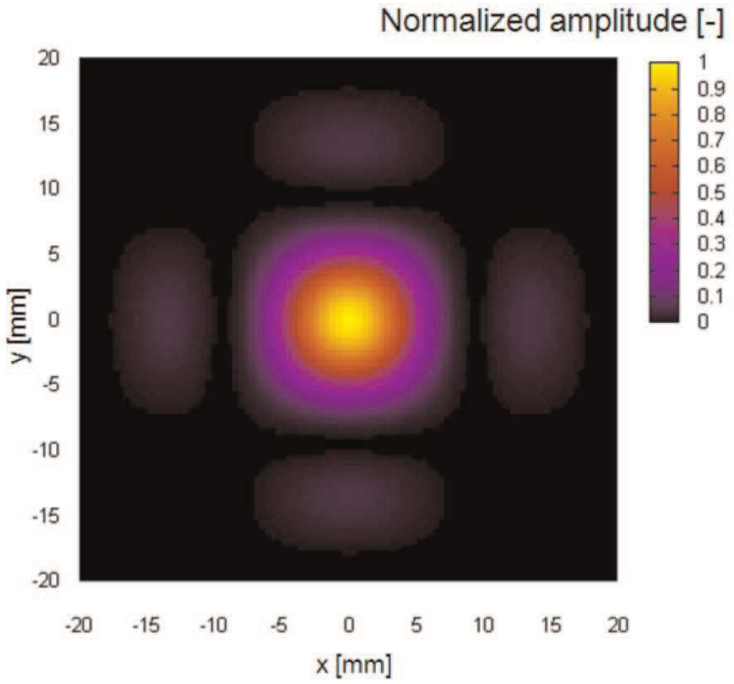

Fig. 3.3. Simulative distribution of focused radiation pressure where $z_f = 200mm$[4]. Ultrasound frequency was set to 40kHz.

AUTD directly limits the depth range of proper focusing. In general, the spot blur size l is given as

$$l = A\frac{\sqrt{D^2 + f^2}}{D}\lambda, \tag{3.25}$$

where A is the constant determined by the condition of image formation and D, f and λ denotes the aperture length, the distance between the focal point and the AUTD surface, and the wavelength respectively. This equation indicates that the spot size remains comparative to the wavelength (diffraction limit) within the depth of the array diameter. Numerical simulation of focusing with two arrays of different aperture length is depicted in Figure 3.4. The wavelength in the simulation was 8.5mm and the focal depth was set to 600mm. It is seen that with an array whose aperture is shorter than the focal depth can not make a focus whose diameter is compatible to the wavelength. It is also confirmed that An array with compatible aperture size creates focusing without blurred.

3.2.2 More Complicated Distribution

Generation of a sound field with more complicated spatial amplitude distribution is discussed in this subsection. It has been known that the spatial Fourier inverse transform of amplitude distribution at the lens is what we have as a projected image on far field. In the previous study, the possibility has already been argued with a concise formulation [4]. There is also another approach to create an image of a planer amplitude distribution in the context of creating two-dimensional tactile image with radiation pressure by adapting the inverse problem framework[41]. What the author describes in the section is similar to [41] yet verifies that it is essentially applicable for producing three-dimensional amplitude distribution with composite arrays of transducers having several transmission planes of multiple directions. The proposed method is also expandable for given additional

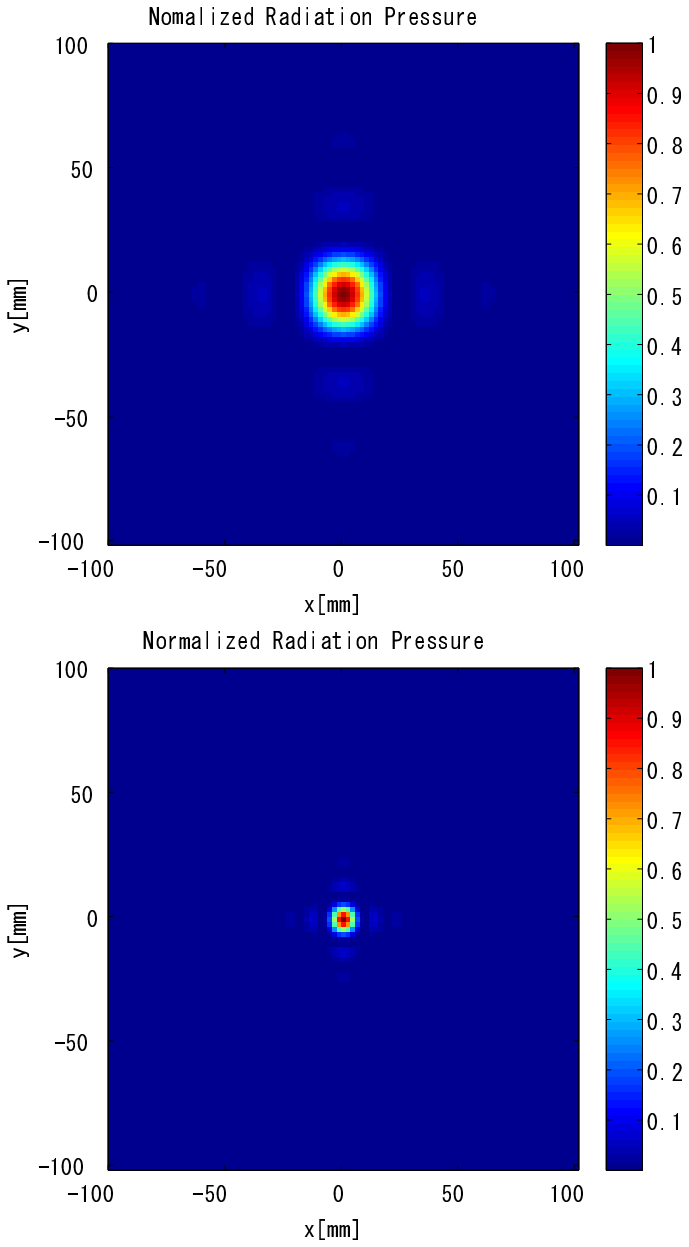

Fig. 3.4. Simulative distribution of focused radiation pressure where $z_f = 600mm$ with an array of 200mm square (Upper Figure) and of 600mm square (Lower Figure). In each transducer array the element interval was set to 10mm. Ultrasound frequency was set to 40kHz.

conditions.

The aim of displaying three-dimensional touchable pressure distribution is to enhance the sense of existence pf a virtual pbject displayed. The author intends to display objects with volume touched with multiple contact points (Figure 3.5). The problem formulation is given in a similar way of a single focus case. The transducer locations are given by $\boldsymbol{r}_i,\ i = 1, \ldots, N$. It is assumed that a pressure distribution which creates the desired three-dimensional tactile form is given as $p(\boldsymbol{q}_l)$ at discrete points $\boldsymbol{q}_l,\ l = 1, \ldots, M$. Let g_i be the complex amplitude of output ultrasound wave emitted from the i-th transducer.

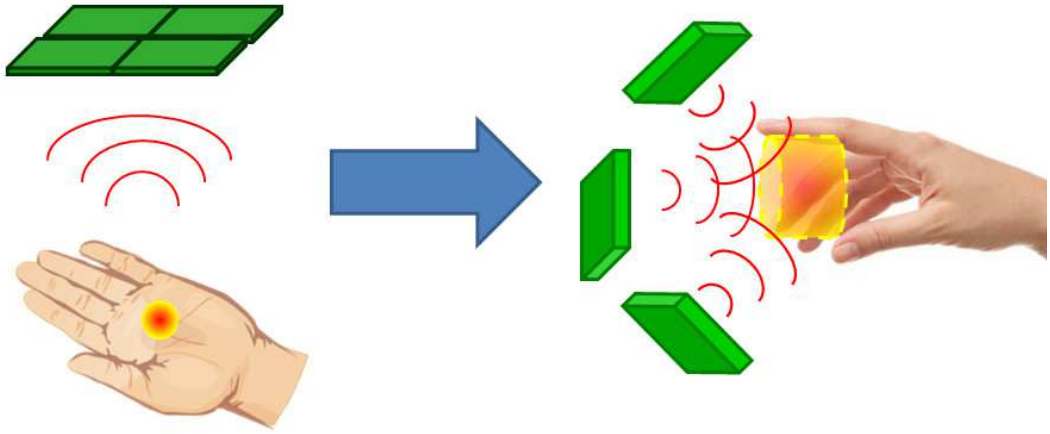

Fig. 3.5. Conceptual image of touchable objects with volume instead of a single tactile focus.

The resulting wave field $p(\boldsymbol{r})$ is obtained by

$$p(\boldsymbol{r}) = \sum_{i=1}^{N} a(\boldsymbol{r}_i, \boldsymbol{r}) e^{-jk|\boldsymbol{r}_i - \boldsymbol{r}|} g_i. \tag{3.26}$$

This is actually almost the same form as (3.23). In the previous problem setting, the only tunable parameter was the phase shift on each transducer. Nevertheless, output waveform from each transducer has two degree of freedom, phase shifts and amplitudes. In this problem formulation, the author describes a method for obtaining the optimal set of output gain g_i with optimal phase shifts and amplitudes. The optimal output set g_i is defined as the one creates a wave field with the highest fidelity compared to given $p(\boldsymbol{q}_l)$.

By introducing an linear-algebraic expression to (3.26), (3.26) can be rewritten as much simpler a form. To begin with, the parameter a_{il} is introduced defined as

$$a_{il} = a(\boldsymbol{r}_i, \boldsymbol{q}_l) e^{-jk|\boldsymbol{r}_i - \boldsymbol{q}_l|}. \tag{3.27}$$

This factor indicates the degree of phase delay and amplitude attenuation of ultrasound emitted from i-th transducer reaching $\boldsymbol{q}_l$ (Figure 3.6). From (3.26), amplitude distribution over $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_M$ is

$$p(\boldsymbol{q}_l) = \sum_{i=1}^{N} a_{il} g_i. \tag{3.28}$$

Next, the following vectors and matrices are defined:

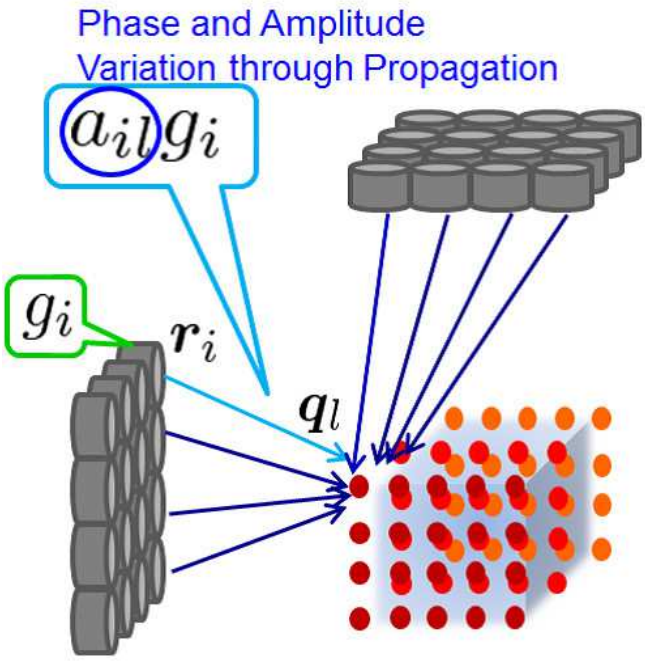

Fig. 3.6. Schematic description of the problem formulation.

$$\boldsymbol{g} := \begin{bmatrix} g_1 \\ \vdots \\ g_N \end{bmatrix}, \tag{3.29}$$

$$\boldsymbol{p} := \begin{bmatrix} p(\boldsymbol{q}_1) \\ \vdots \\ p(\boldsymbol{q}_M) \end{bmatrix}, \tag{3.30}$$

$$\boldsymbol{A} := \begin{bmatrix} a_{11} & \cdots & a_{N1} \\ \vdots & \ddots & \vdots \\ a_{1M} & \cdots & a_{NM} \end{bmatrix}. \tag{3.31}$$

Using these parameters, (3.28) is rewritten as

$$\boldsymbol{p} = A\boldsymbol{g}. \tag{3.32}$$

Now that $\boldsymbol{p}$ is already determined because every component $p(\boldsymbol{q}_l)$ is given, the optimal output amplitude vector $\boldsymbol{g}$ is what to obtain. By introducing a proper criterion to evaluate the degree of optimization, the optimal $\boldsymbol{g}$ is obtained as the one which optimizes the evaluation function. In this kind of optimization, the squared error is often chosen as the evaluation function to be minimized:

$$\begin{aligned} J &= |\boldsymbol{p} - A\boldsymbol{g}|^2 \\ &= (\boldsymbol{p} - A\boldsymbol{g})^H(\boldsymbol{p} - A\boldsymbol{g}) \\ &= \boldsymbol{p}^H\boldsymbol{p} - \boldsymbol{p}^H A\boldsymbol{g} - \boldsymbol{g}^H A^H\boldsymbol{p} + \boldsymbol{g}^H A^H A\boldsymbol{g}. \end{aligned} \tag{3.33}$$

Here $\cdot^H$ denotes the conjugate transpose. One of general minimization procedure is

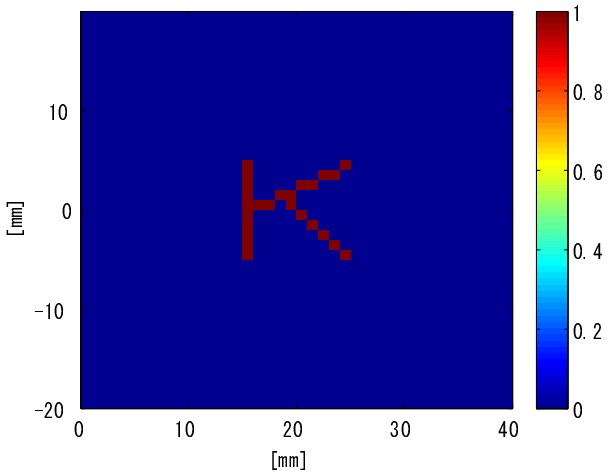

Fig. 3.7. Ideal normalized spatial amplitude distribution of generated sound pressure

demonstrated in the following. Fisrst, obtain the partial derivative of J by $\boldsymbol{g}^H$:

$$\frac{\partial J}{\partial \boldsymbol{g}^H} = A^H \boldsymbol{p} - A^H A \boldsymbol{g}. \tag{3.34}$$

At the minimal point, the derivative has to vanish:

$$A^H \boldsymbol{p} - A^H A \boldsymbol{g} = 0. \tag{3.35}$$

Hence, the optimal $\boldsymbol{g}$ is given as

$$\boldsymbol{g} = (A^H A)^{-1} A^H \boldsymbol{p}. \tag{3.36}$$

Note that the same outcome is obtained by differentiating (3.33) with $\boldsymbol{g}$ instead of $\boldsymbol{g}^H$.

Numerical Experiments

We validated the effectiveness of our method through some numerical simulative experiments. Figure 3.7 shows the desired spatial distribution of sound pressure amplitude. Note that although in the experiments the produced pattern was defined only in two-dimensional plane, the proposed method can be applied to the patterns spreading in three-dimensional way. The pattern was defined at discrete points in lattice for every 10 mm. The simulation for generating this pattern was conducted for two different positions of pattern generating plane (1, 2) and two different arrangements of transducers (a, b). So we have results of simulations under four different combinations of each two experimental condition, (1a), (1b), (2a), (2b). These conditions are schematically depicted in Figure 3.8. The region (1) is a square xy-plane of 400 mm x 400mm and centered on the z axis at z = 200mm. The other region (2) is a square of same size which is parallel to the yz-plane and centered on the coordinate (x, y, z) = (0, 0, 200mm). The transducer arrangement (a) is a single planar lattice of every 10 mm. The lattice region is a part of the xy-plane centered on the coordinate origin and its size is 20mm in x axis and 40mm in y axis. The other arrangement (b) is created by folding the arrangement (a) in half toward the direction of the z-axis so that the ridge of the resultant transducer array may remain in contact with the xy-plane. For each of these four arrangements, the optimum output gain of each transducer was calculated by (3.36) and simulated the spatial distribution of

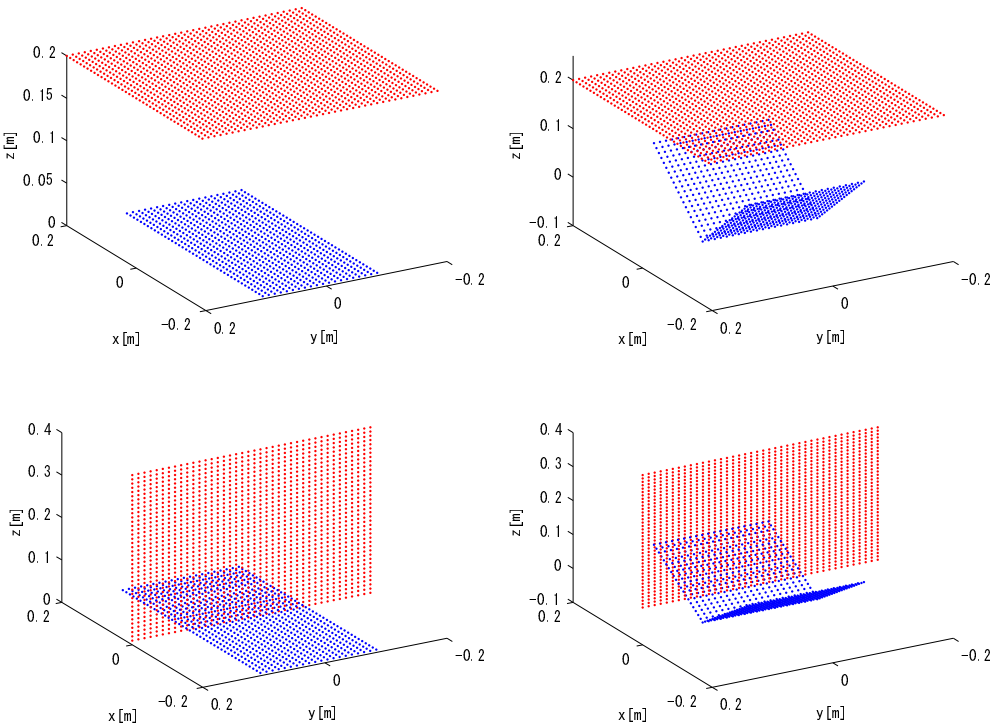

Fig. 3.8. Four location sets of transducers(blue points) and regions in which spatial patterns of sound pressure are generated(red points), 1a(top), 1b(second), 2a(third) and 2b(bottom) depicted respectively

amplitude produced by the obtained output sound waves. In this numerical experiment, the attenuate coefficients a_{il} was defined as

$$a_{il} = \frac{1}{|\boldsymbol{r}_i - \boldsymbol{q}_l|} \tag{3.37}$$

The results are shown in Figure 3.9. Little difference in the resultant spatial amplitude between (1a) and (1b) is seen. On the contrary it is obvious that the imaging performance in (2a) is more degraded than in (2b). The spatial detail and maximum power are both lower in (2a). The reason is because it is difficult for transducers to make sharp focuses in the depth (z-axis) direction due to small extent of transducer locations in the xy-plane and no spatial flexibility of arrangement in z-axis. In the case of the arrangement (2b), the positional freedom of transducers is guaranteed in all of x, y and z axis. The experiment has shown that the spatial freedom in transducer arrangement directly affects the imaging performance. The directive dependency of imaging performance can be unbiased by spatially uniform transducer arrangement. Note that this simulation was done with MATLAB numerical computing engines. In the simulation the solution of 3.35 was obtained with MATLAB functions. Since the case of (2a) and (2b) the problem had the ill-posed property, the obtained solution is based on other criteria such as minimizing the norm of the solution.

The output amplitudes of transducers are preferable to be close to their possible maximum as much as possible. The lower part of figure 3.9 shows the histogram of output amplitude obtained by the optimization. For each condition, it is seen that only a small number of transducers contributes to image production.

It is considered that this is because most transducers increase the error function with

any phase shift in given conditions. Even if it is true, since tactually perceivable sound amplitude is really intense, small excessive output at the region where no sound should be created would be negligible. By constructing better evaluation functions which take priority over the intense output power instead of complete null at desired region, the number of contributing transducer would be increased.

3.2.3 Modification of Evaluation Functions

The evaluation function derived at the previous subsection can be considered so straightforward that it does not yield preferable pressure distributions under some conditions. As stated above, tactually perceivable sound amplitude is above 150 dB$\sim$, which means unintended output whose intensity is less than that does not violate the designed output tactile stimuli. On the other hand, it is expected that deficiency of output radiation pressure at the desired region directly deteriorates generated tactile stimuli. In short, in the optimization of pressure distributions, output errors at regions where intense output pressure is expected should be sacrificed in order not to reduce the output intensities in desired regions.

Based on this idea, a new evaluation function to optimize is given as the following:

$$J(\boldsymbol{g}) := \sum_{l=1}^{N} w_l \left| \sum_{i=1}^{M} a_{il} g_i - p_l \right|^2 , \tag{3.38}$$

where w_l denotes the weighting coefficients corresponding to the given output pressure p_l. These coefficients determine how the errors at $\boldsymbol{q}_l$ are critical for the optimization. If w_l with a certain value of l is higher than that of other regions, errors at bmq_l has more influences on the resulting value of the evaluation functions. The most extreme case is that the values of w_l are 0 for some l, which means the output pressure value p_l at $\boldsymbol{q}_l$ causes no variance of the evaluation functions. Therefore it is equivalent to that no errors in these regions are taken into account. In the current context, we want to derive evaluation functions which take priority over reducing the errors at the region where the output pressure should be intense. Hence the weighting coefficients are supposed to be a function of the given desired output pressure p_l, such as $w_l = w(\boldsymbol{p}_l)$. The weighting coefficients should increase as the value of p_l increases so that the value of the evaluation function becomes sensitive to the output errors in regions where high pressure is supposed to be generated. Here are some examples:

$$w(p) = (|p| + \epsilon), \quad w(p) = e^{a|p|}. \tag{3.39}$$

The term ϵ is introduced in order to prevent the output errors at $\boldsymbol{q}_l$ where given p_l are 0 from being completely neglected. The value of ϵ determines how much output errors are to be negligible where p_l has relatively small values.

The modified function (3.38) can be rewritten as

$$J(\boldsymbol{g}) = |W(A\boldsymbol{g} - \boldsymbol{p})|^2 , \tag{3.40}$$

where the weighting matrix W is given by the following diagonal matrix:

$$W := \begin{bmatrix} \sqrt{w_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{w_M} \end{bmatrix} . \tag{3.41}$$

By the similar procedure used in deriving the optimal solutions of (3.33), the optimal value of $\boldsymbol{g}$ is given by

$$\boldsymbol{g} = (A^H W^H W A)^{-1} A^H W^H W \boldsymbol{p}. \tag{3.42}$$

This solution is obtained by substituting WA instead of A and $W\boldsymbol{p}$ instead of $\boldsymbol{p}$ in (3.36).

Experimental results shown in the previous subsection implies that only a small number of transducers contribute to producing given pressure pattern. There is a problem for this result. Transducers with high output intensity is required to reconstruct the obtained optimal pressure distribution in practice. This fact means that if the required value is excessive for the actual It is true that there is a limit determined by the transducer power, in some cases this limit can be overcome by adjusting the output intensity unbiased among the transducers. For this reason, it is preferable that output amplitude of each transducer are equalized. This is equivalent to minimize the variance of g_i. It is expressed as

$$E\left[|g_i|^2\right] - \left(E\left[g_i\right]\right)^2, \tag{3.43}$$

where $E[\cdot]$ denotes the averaging operator. (3.43) can be rewritten as

$$\frac{1}{N}\boldsymbol{g}^H\boldsymbol{g} - \left|\frac{1}{N}\mathbf{1}^H\boldsymbol{g}\right|^2, \tag{3.44}$$

where vector $\mathbf{1} := [1 \ \ldots \ 1]^T$ is a column vector of length N. (3.43) can be rewritten as

$$\frac{1}{N}\boldsymbol{g}^H\boldsymbol{g} - \frac{1}{N^2}\boldsymbol{g}^H\mathbf{1}\mathbf{1}^H\boldsymbol{g} = \boldsymbol{g}^H\left(\frac{1}{N}I - \frac{1}{N^2}\mathbf{1}\mathbf{1}^H\right)\boldsymbol{g}, \tag{3.45}$$

where I denotes the unit matrix. Note that the matrix $\mathbf{1}\mathbf{1}^H$ is given as

$$\mathbf{1}\mathbf{1}^H = \begin{bmatrix} 1 & \ldots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \ldots & 1 \end{bmatrix}. \tag{3.46}$$

By defining

$$V := \left(\frac{1}{N}I - \frac{1}{N^2}\mathbf{1}\mathbf{1}^H\right), \tag{3.47}$$

the variance (3.45)is eventually rewritten as

$$\boldsymbol{g}^H V \boldsymbol{g}. \tag{3.48}$$

Another modified evaluation function is constructed by adding this term to (3.38) as another quantity to be minimized:

$$J(\boldsymbol{g}) = |W(A\boldsymbol{g} - \boldsymbol{p})|^2 + \gamma \boldsymbol{g}^H V \boldsymbol{g}, \tag{3.49}$$

where γ denotes the coefficient determining the degree of influences over the total optimization of the evaluation function. A larger γ means that less variance of g_i is tolerable. The partial derivative of (3.49) by $\boldsymbol{g}$ equals to a zero vector if and only if

$$(A^H W^H W A + \gamma V)\boldsymbol{g} = A^H W^H W \boldsymbol{p}. \tag{3.50}$$

Finally, the modified optimal solution is

$$\boldsymbol{g} = (A^H W^H W A + \gamma V)^{-1} A^H W^H W \boldsymbol{p}. \tag{3.51}$$

Since the maximum output amplitude among all transducers limits the allowable output amplitude scale, this minimizing-variance strategy is expected to enable to generate radiation pressure distribution with reinforced intensity.

There is another possible strategy for making use of possible maximum output power of transducers, in which only phase of each transducer is allowed to vary. In this case the method of Lagrange multiplier is applicable. The modified objective function is given by

$$J_{\boldsymbol{\lambda}} = |\boldsymbol{p} - A\boldsymbol{g}|^2 - \sum_{i=1}^{N} \lambda_i (g_0 - |g_i|), \tag{3.52}$$

where g_0 denotes the common amplitudes of transducers, with N Lagrange multiplier $\lambda_1, \ldots, \lambda_N$. The thesis does not derive the detailed solution of (3.52).

Numerical Experiments

We operated numerical experiments for evaluating the performances of those modified evaluation functions. All experimental conditions except the modifications on the evaluation function (3.49) were the same as the experiment in the last subsection. The experiments conducted are divided into three parts:

- Experiment 1: Evaluation of weighting matrix, where

$$w_l = |p_l| + \epsilon, \ \epsilon = 0.001, \ \gamma = 0.$$

- Experiment 2: Evaluation of minimization of output amplitude variances, where

$$W = I, \ \gamma = 100.$$

- Experiment 3: Evaluation of both strategies above with a set of ϵ and γ, where

$$w_l = |p_l| + \epsilon, \ \epsilon \in \{1, 0.001, 0.00001\}, \ \gamma \in \{1, 100, 100000\}.$$

Figure (3.10) shows the result for the Experiment 1. For all spatial arrangements, output amplitude is seen magnified where given amplitude is relatively intense. In those regions, higher fidelity in output amplitude is achieved compared to the previous optimizations. On the other hand, unintended energy emittance in where there should be no output intensity is observed at higher levels than those in the previous experiment. Another significant difference is the highest value in obtained output amplitude among transducers, especially in the case (2a) and (2b). Obtained sets of transducer output amplitude with highly biased values are practically infeasible because they require extraordinary dynamic output ranges.

In the experiment 2, effects of the variance averaging term $\gamma \boldsymbol{g}^H V \boldsymbol{g}$ is observed (3.11). It is seen the variance of transducer output amplitude is drastically reduced for all spatial conditions preventing output pressure distribution from being deteriorated compared to the first experiment in the previous subsection 3.9. This result suggests that adding the variance averaging terms can enable to generate desired pattern by transducers with smaller output intensity. Thus this modification leads to magnifying output levels.

The summary of the two experiments above are:

1. The weighting matrix W creates output pressure distribution with sufficient intensity, while extremely high transducer amplitude is occasionally required.
2. Adding the variance term $\gamma \boldsymbol{g}^H V \boldsymbol{g}$ inhibits the variance of output amplitude being too large.

The derived evaluation function (3.49) is a linear combination of two terms with tunable parameter ϵ and γ. In the experiment 3, we operated numerical experiments for the arrangement (2b) with 9 combinations of ϵ and γ (Figure 3.12). It is ascertained that more priority over output amplitude variance minimization is taken as γ increases, while greater output pressure fidelity at where the given pattern has high value is observed under smaller ϵ. As ϵ decreases, errors in regions in which relatively small output amplitude is supposed to be produced have less influence on the optimization. As a result, those errors are neglected. The result of this experiment shows that an appropriate set of ϵ and γ can make the function of the matrix W and the term $\gamma \boldsymbol{g}^H V \boldsymbol{g}$ compatible. To sum up, sufficient pressure distribution can be obtained with averaged output amplitude of each transducer.

Note that the rank of matrix V is $N-1$. It means that for sufficiently large γ, the matrix $(A^H W^H W A + \gamma V)$ becomes non-singular. In the case, no additional constraint like norm-minimization of obtained solution is unnecessary and it is guaranteed that the inverse matrix of $(A^H W^H W A + \gamma V)$ exists. Thie term $\gamma \boldsymbol{g}^H V \boldsymbol{g}$ can be regarded as a regularization term in optimization.

3.3 Particle Velocity Control

In the modeling above, the direction of energy propagation is not taken into account. For the proper generation of tactual images with volume, direction control of radiation pressure should be an important issue. The energy flow direction is determined by the sound field. What determines the direction of energy flux is the particle velocity vector $\boldsymbol{v}$. As stated above, the particle velocity distribution $\boldsymbol{v}(\boldsymbol{r}, t)$ is obtained as the spatial gradient of acoustic potential $\phi(\boldsymbol{r}, t)$. Let $\boldsymbol{r}_i,\ i = 1, \ldots, N$ be the positions of transducers and $p(\boldsymbol{r}, t)$ be the instantaneous pressure at $\boldsymbol{r}$. The relationships between these three physical quantities are:

$$p(\boldsymbol{r}, t) = \rho \frac{\partial \phi}{\partial t} \tag{3.53}$$

$$\boldsymbol{v}(\boldsymbol{r}, t) = -\nabla \phi(\boldsymbol{r}, t). \tag{3.54}$$

The resulting wave field by given transducer output is completely the same as (3.26):

$$p(\boldsymbol{r}) = \sum_{i=1}^{N} b(\boldsymbol{r}_i, \boldsymbol{r}) g_i. \tag{3.55}$$

where $b(\boldsymbol{r}_i, \boldsymbol{r}) := a(\boldsymbol{r}_i, \boldsymbol{r}) e^{-jk|\boldsymbol{r}_i - \boldsymbol{r}|}$. By assuming time-harmonic fields, time differential of those quantities are simple given by multiplying them by $j\omega$. From (3.55),(3.53) and (3.54),

$$\phi(\boldsymbol{r}, t) = \frac{1}{j\omega\rho} p(\boldsymbol{r}, t) \tag{3.56}$$

Therefore,

$$\boldsymbol{v}(\boldsymbol{r}, t) = \frac{j}{\omega\rho} \nabla p(\boldsymbol{r}, t) \tag{3.57}$$

$$= \frac{j}{\omega\rho} \nabla \sum_{i=1}^{N} g_i b(\boldsymbol{r}_i, \boldsymbol{r}) e^{\omega t} \tag{3.58}$$

$$= \frac{j}{\omega\rho} e^{j\omega t} \sum_{i=1}^{N} g_i \nabla_{\boldsymbol{r}} b(\boldsymbol{r}_i, \boldsymbol{r}). \tag{3.59}$$

By removing the time variant term, we have

$$\boldsymbol{v}(\boldsymbol{r}) = \frac{j}{\omega\rho}\sum_{i=1}^{N} g_i \nabla b_{\boldsymbol{r}}(\boldsymbol{r}_i, \boldsymbol{r}). \tag{3.60}$$

The equation (3.60) indicates that the direction of the particle velocity is obtained by the sum of the gradient vector $\nabla b_{\boldsymbol{r}}(\boldsymbol{r}_i, \boldsymbol{r})$ weighted by the coefficient g_i. Since $\nabla b_{\boldsymbol{q}}(\boldsymbol{r}_i, \boldsymbol{q})$ is determined once the propagation model function $a(\boldsymbol{r}_i, \boldsymbol{r})$, every source position $\boldsymbol{r}_i$ and the location of the interest $\boldsymbol{q}_l,\ l = 1, \ldots, N$ are given, the only tunable parameters are source amplitudes g_i as the previous problem formulation.

Resultant particle velocity field is derived from a given set of source positions, source amplitudes and their phase delays. From here the author derives the optimal set of them from the given particle velocity distribution by applying the inverse problem method. As the case in the previous session, the desirable particle velocity vectors are assumed to be given. How to create the optimal velocity set to generate a designed tactile object currently remains as a future work in the research.

To handle the problem more simplistic way, we again introduce a linear algebraic representation. Since the gradient vectors $\nabla b_{\boldsymbol{q}}(\boldsymbol{r}_i, \boldsymbol{q})$ can be determined once the model function is chosen, we have a set of vector $\boldsymbol{d}_i(\boldsymbol{q})$ corresponding to the source position $\boldsymbol{r}_i$:

$$\boldsymbol{d}_i(\boldsymbol{q}) := \nabla b_{\boldsymbol{q}}(\boldsymbol{r}_i, \boldsymbol{q}). \tag{3.61}$$

By using these expressions, the resultant particle vector has the complex form:

$$\boldsymbol{v}(\boldsymbol{q}) = \frac{j}{\omega\rho}\sum_{i=1}^{N} g_i \boldsymbol{d}_i(\boldsymbol{q}). \tag{3.62}$$

Next, let $\tilde{\boldsymbol{v}}(\boldsymbol{r})$ be the desired distribution of particle velocity determined in a region Ω, we have Euclidean norm of error vectors:

$$|\tilde{\boldsymbol{v}}(\boldsymbol{q}) - \boldsymbol{v}(\boldsymbol{q})|^2. \tag{3.63}$$

The summed error is:

$$\int_{\Omega} |\tilde{\boldsymbol{v}}(\boldsymbol{q}) - \boldsymbol{v}(\boldsymbol{q})|^2 dV = \int_{\Omega} |\tilde{\boldsymbol{v}}(\boldsymbol{q}) - \frac{j}{\omega\rho}\sum_{i=1}^{N} g_i \boldsymbol{d}_i(\boldsymbol{q})|^2 dV. \tag{3.64}$$

It is expected that we can create the desired particle velocity distribution $\tilde{\boldsymbol{v}}$ by minimizing Eq.(3.64) by the optimal value set of $g_i,\ i = 1, \ldots, N$. A straightforward way for obtaining the optimal solution of Eq.(3.64) is to obtain partial derivative of (3.64) for every g_i and find a parameter g_i which vanishes them. In case of the summed error has a form of integral, the partial derivatives of it varies depending on the region Ω and the desired particle velocity distribution $\boldsymbol{v}(\boldsymbol{q})$. Since a common procedure for optimizing them is very hard to find and to describe a few example of them will not help designing a wide variety of tangible objects, in the following we deal with a case with discreet velocity distribution $\boldsymbol{v}(\boldsymbol{q}_l)$ and region $\Omega_d := \{\boldsymbol{q}_l |\ \boldsymbol{q}_1, \ldots, \boldsymbol{q}_M\}$. Then a new error function can be defined:

$$\sum_{l=1}^{M} |\tilde{\boldsymbol{v}}(\boldsymbol{q}_l) - \frac{j}{\omega\rho}\sum_{i=1}^{N} g_i \boldsymbol{d}_i(\boldsymbol{q}_l)|^2. \tag{3.65}$$

Here we introduce some vectors and matrices again in order to simplify the error function above. Let $\boldsymbol{g}$, $\boldsymbol{d}_{il}, \boldsymbol{v}_l$ and B_l as following:

$$\boldsymbol{g} := \begin{bmatrix} g_1 \\ \vdots \\ g_N \end{bmatrix}, \tag{3.66}$$

$$\boldsymbol{d}_{il} := \boldsymbol{d}_i(\boldsymbol{q}_l), \tag{3.67}$$

$$\tilde{\boldsymbol{v}}_l := \tilde{\boldsymbol{v}}(\boldsymbol{q}), \tag{3.68}$$

$$B_l := j\frac{1}{\omega\rho}[\boldsymbol{d}_{1l} \ \cdots \ \boldsymbol{d}_{Nl}]. \tag{3.69}$$

By these expressions, (3.65) can be rewritten as

$$J(\boldsymbol{g}) = \sum_{l=1}^{M} |\tilde{\boldsymbol{v}}_l - B_l\boldsymbol{g}|^2. \tag{3.70}$$

The optimal $\boldsymbol{g}$ is given by

$$\boldsymbol{g} = \left(\sum_{l=1}^{M} B_l^H B_l\right)^{-1} \left(\sum_{l=1}^{M} B_l^H \tilde{\boldsymbol{v}}_l\right). \tag{3.71}$$

In the previous subsection, it has been demonstrated that the modifications of evaluation functions is effective for obtaining practical solutions. Here the same procedure can be applied. The weighting coefficients w_l can be defined as a function of given particle velocity $\tilde{\boldsymbol{v}}_l$. By adding the variance term, (3.70) is modified as

$$J(\boldsymbol{g}) = \sum_{l=1}^{M} w_l|\tilde{\boldsymbol{v}}_l - B_l\boldsymbol{g}|^2 + \gamma\boldsymbol{g}^H V\boldsymbol{g}. \tag{3.72}$$

For this fuction, the optimal $\boldsymbol{g}$ is given by

$$\boldsymbol{g} = \left(\sum_{l=1}^{M} w_l B_l^H B_l + \gamma V\right)^{-1} \left(\sum_{l=1}^{M} w_l B_l^H \tilde{\boldsymbol{v}}_l\right). \tag{3.73}$$

Numerical Experiments

For the experiment, the tuning parameters were set to $\gamma = 100000, \epsilon = 0.01$, where $w_l = |\tilde{\boldsymbol{v}}_l| + \epsilon$. In the experiment, we intended to generate a pattern of planer velocity distributions. For simplicity, the propagation model function $a(\boldsymbol{r}_i, \boldsymbol{q})$ was assumed to be constant value 1. Under this assumption, the term $b_{\boldsymbol{q}}(\boldsymbol{r}_i, \boldsymbol{q})$ becomes

$$b_{\boldsymbol{q}}(\boldsymbol{r}_i, \boldsymbol{q}) = e^{-jk|\boldsymbol{r}_i - \boldsymbol{q}|}. \tag{3.74}$$

Eventually, its gradient vector can be represented as

$$\nabla b_{\boldsymbol{q}}(\boldsymbol{r}_i, \boldsymbol{q}) = \frac{jk}{|\boldsymbol{r}_i - \boldsymbol{q}|} e^{-jk|\boldsymbol{r}_i - \boldsymbol{q}|}(\boldsymbol{r}_i - \boldsymbol{q}). \tag{3.75}$$

This result indicates that the particle velocity for the given transducer output $\boldsymbol{g}$ is obtained as

$$\boldsymbol{v}(\boldsymbol{q}) = -\frac{k}{\omega\rho} \sum_{i=1}^{N} g_i \frac{1}{|\boldsymbol{r}_i - \boldsymbol{q}|} e^{-jk|\boldsymbol{r}_i - \boldsymbol{q}|}(\boldsymbol{r}_i - \boldsymbol{q}). \tag{3.76}$$

This result is intuitively understandable in terms of the fact that direction of the particle velocity is parallel to $(\boldsymbol{q} - \boldsymbol{r}_i)$, the direction from transducers to the corresponding positions. The reconstruction of sound field from given source amplitude $\boldsymbol{g}$ in this experiment is done according to (3.76).

We numerically reconstructed two planar patterns of particle velocity distribution (Figure3.13, Figure3.14). Note that the obtained particle velocity $\boldsymbol{v}(\boldsymbol{r})$ is a complex vector since it has three phases with respect to its coordinate axis. The imaginary part of obtained particle velocity is depicted in the figure. In the figure, obtained sound intensity vectors $\boldsymbol{I}(\boldsymbol{q})$ for spatial location $\boldsymbol{q}$, which indicate the direction of energy flux at $\boldsymbol{q}$, is depicted:

$$\boldsymbol{I}(\boldsymbol{q}) = \mathrm{Re}[p(\boldsymbol{q})\boldsymbol{v}^*(\boldsymbol{q})], \tag{3.77}$$

where $\mathrm{Re}[\cdot]$ denotes the real part of $\cdot$ and $\cdot^*$ denotes the complex conjugate of $\cdot$ respectively. Each component of $\boldsymbol{I}(\boldsymbol{q})$ is represented as

$$|p||v_i| \cos \phi_i, \tag{3.78}$$

where v_i is i-th component of $\boldsymbol{v}$, v_x, v_y, v_z, namely. ϕ_i is a phase difference between p and v_i.

In both experimental results, it is observed that the direction of reconstructed particle velocity is almost parallel to the given ones at regions where given particle velocity has nonzero value. However their surrounding areas also contain accompanying particle velocity, especially in Fig.3.14. It is significant that errors in sound intensity domain are smaller than those in velocity domain although the optimization is only done in the particle velocity domains. As to the sound pressure, it is seen that it was less localized when compared to those in the previous subsection. There is a possibility that a more adroit optimization such as that in energy flux domain can improve these results.

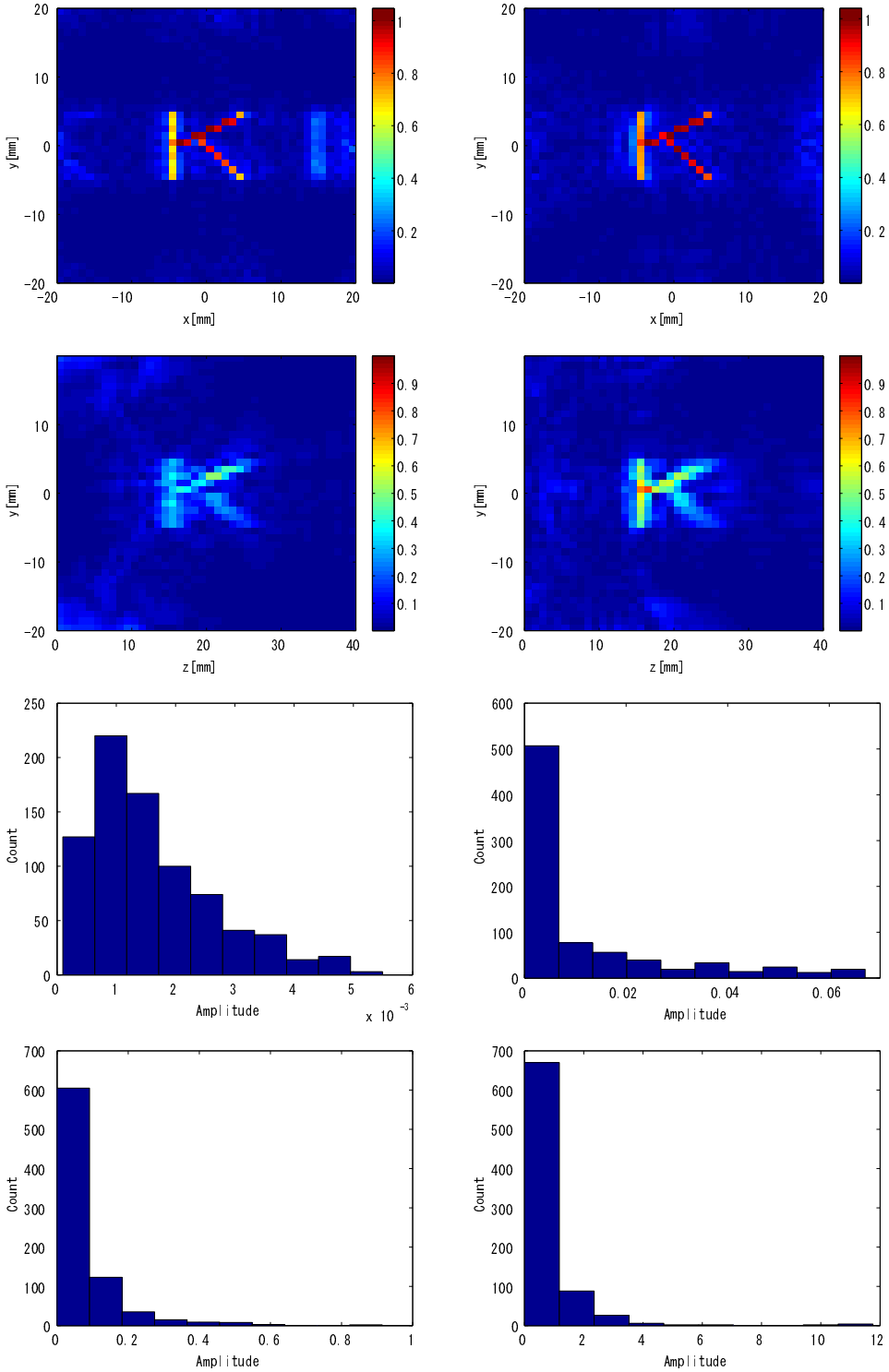

Fig. 3.9. Normalized spatial amplitude distribution of sound pressure generated for the location 1a(top left),1b(top right), 2a(2nd row left), 2b(2nd row right)and histogram of transducer amplitudes for corresponding location conditions 1a(3rd row left),1b(3rd row right), 2a(bottom left), 2b(bottom right) obtained by the optimization of (3.33).

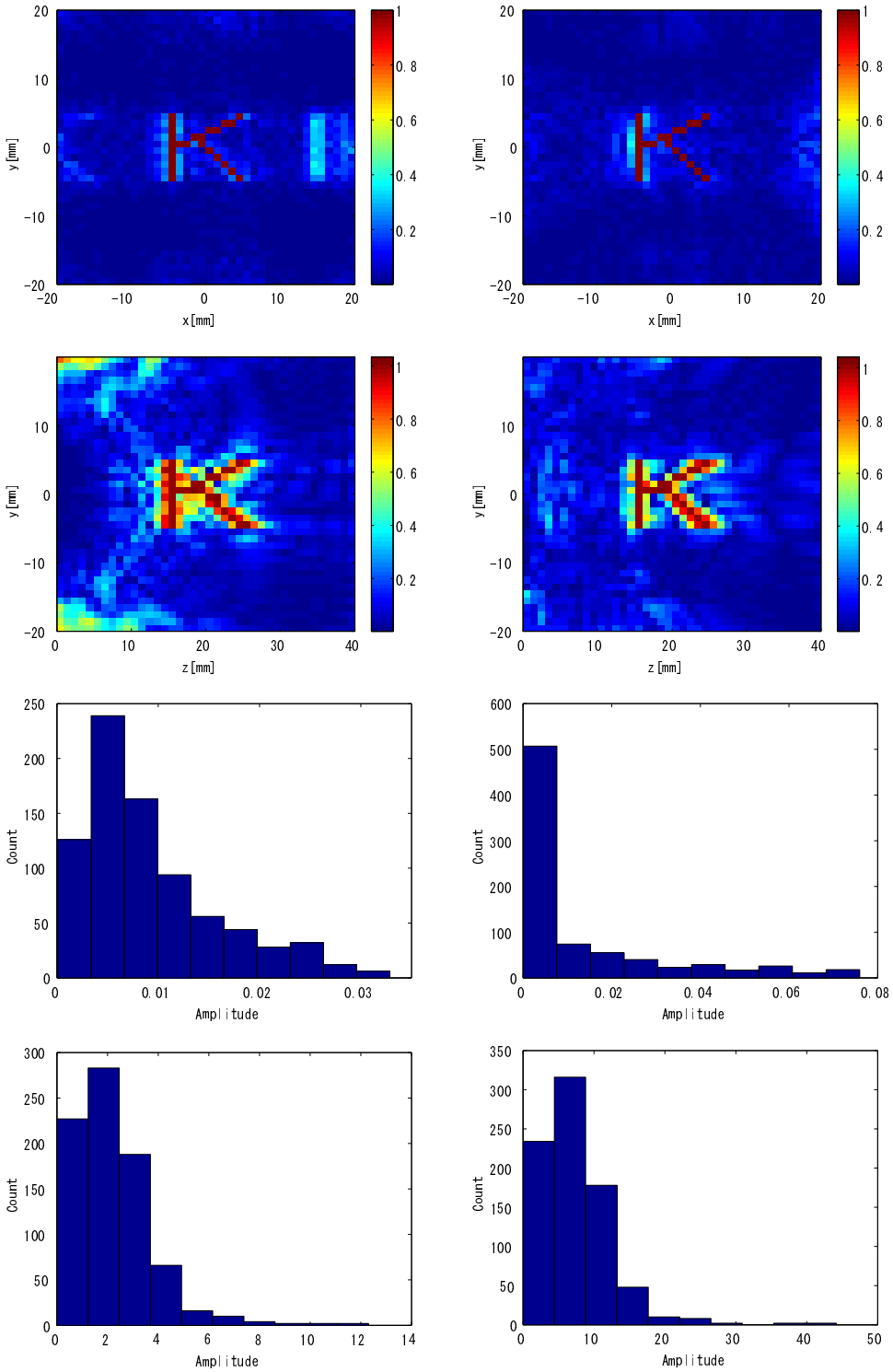

Fig. 3.10. Normalized spatial amplitude distribution of sound pressure generated for the location 1a(top left),1b(top right), 2a(2nd row left), 2b(2nd row right)and histogram of transducer amplitudes for corresponding location conditions 1a(3rd row left),1b(3rd row right), 2a(bottom left), 2b(bottom right) obtained by the optimization of (3.49) in experiment 1.

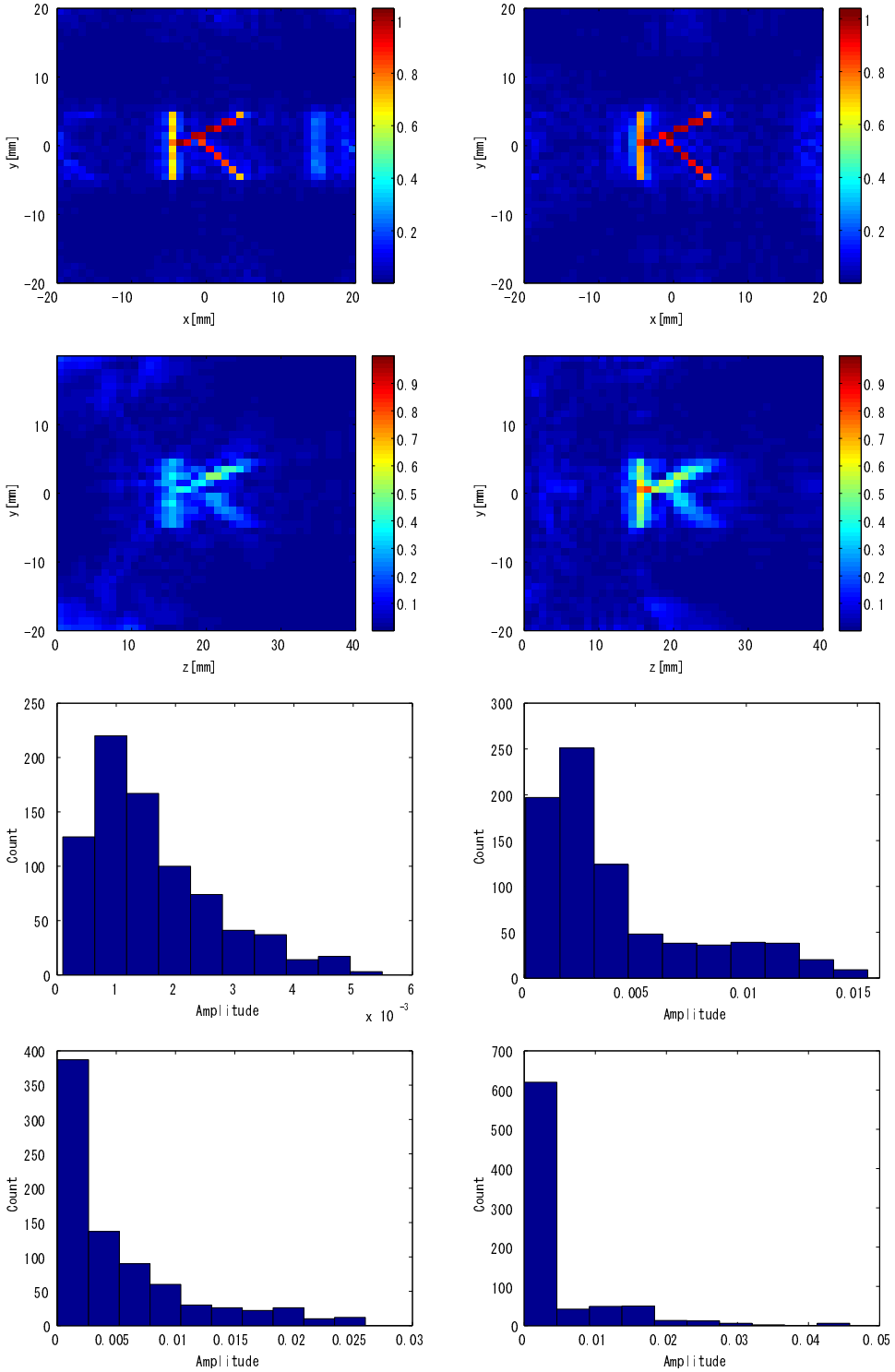

Fig. 3.11. Normalized spatial amplitude distribution of sound pressure generated for the location 1a(top left),1b(top right), 2a(2nd row left), 2b(2nd row right)and histogram of transducer amplitudes for corresponding location conditions 1a(3rd row left),1b(3rd row right), 2a(bottom left), 2b(bottom right) obtained by the optimization of (3.49) in experiment 2.

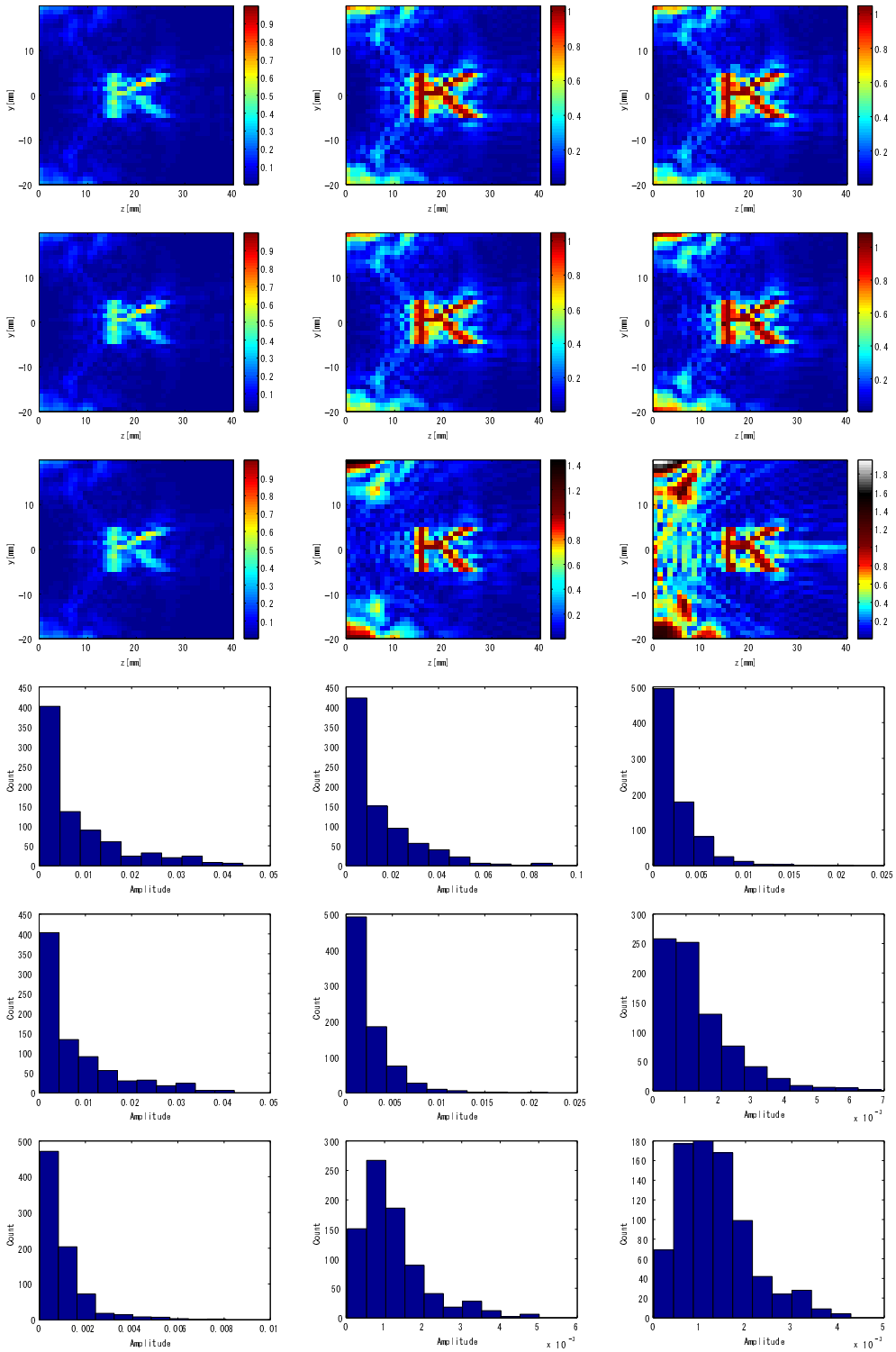

Fig. 3.12. Normalized spatial amplitude distribution of sound pressure generated for the location (2a) with different combinations of parameter values, $\gamma = 1$(1st row), $\gamma = 100$(2nd row), $\gamma = 100000$(3rd row), $\epsilon = 1$(left column), $\epsilon = 0.001$(center column) and $\epsilon = 0.00001$(right column) and histogram of transducer amplitudes for corresponding parameter combinations (from 4th to the bottom rows) in experiment 3.

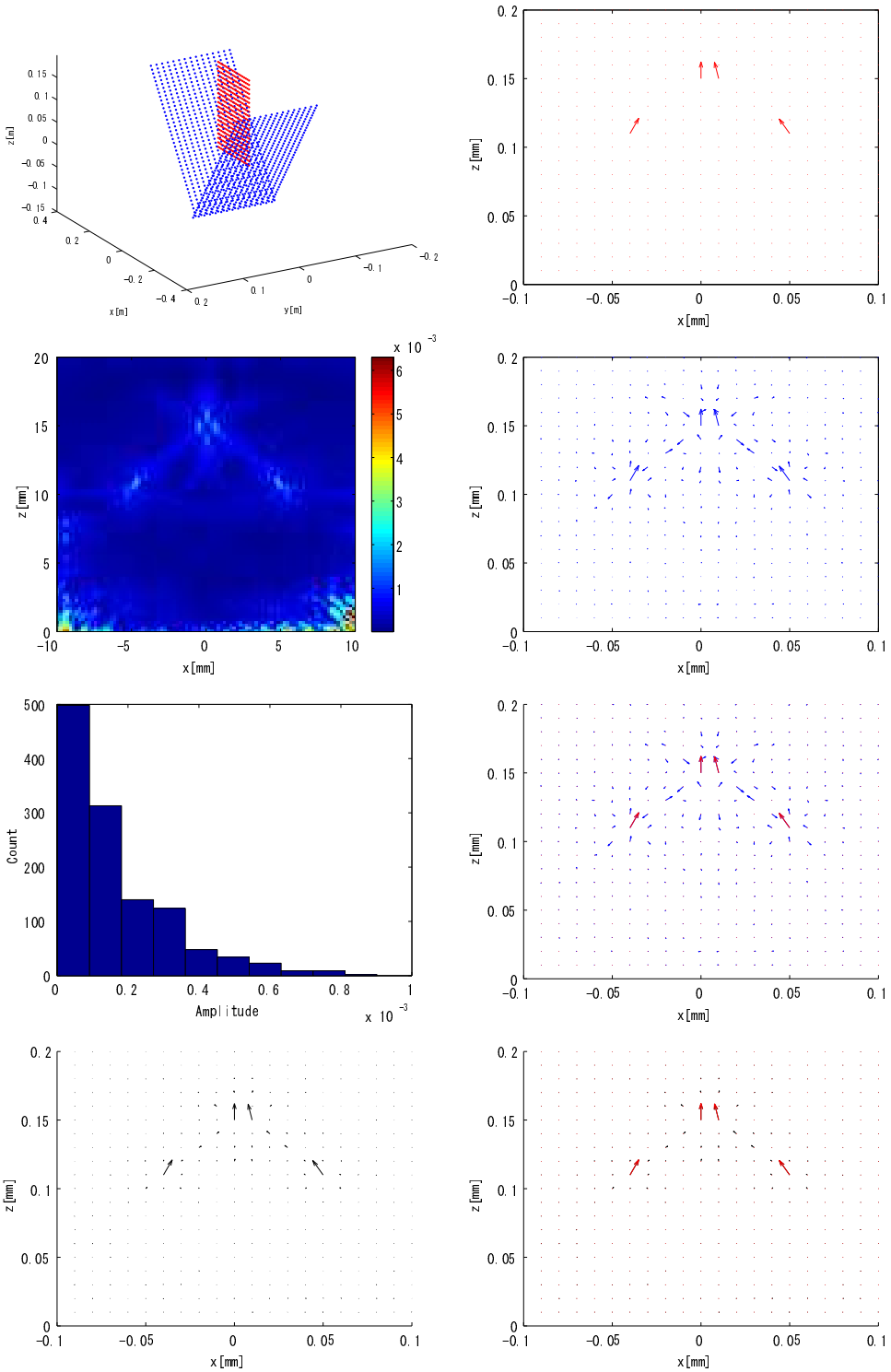

Fig. 3.13. Numerical experimental results for given location condition (left top, blue dots indicate transducers and red dots indicate where the particle velocity is given) and given particle velocity distribution (right top). The obtained absolute amplitude distribution (2nd row left), imaginary part of velocity distribution (2nd row right), output amplitude histogram (3rd row left). Given and obtained velocity distribution are shown overlapped (3rd row right). Obtained acoustic intensity (bottom left), overlapped by particle velocity (bottom right). All values are normalized.

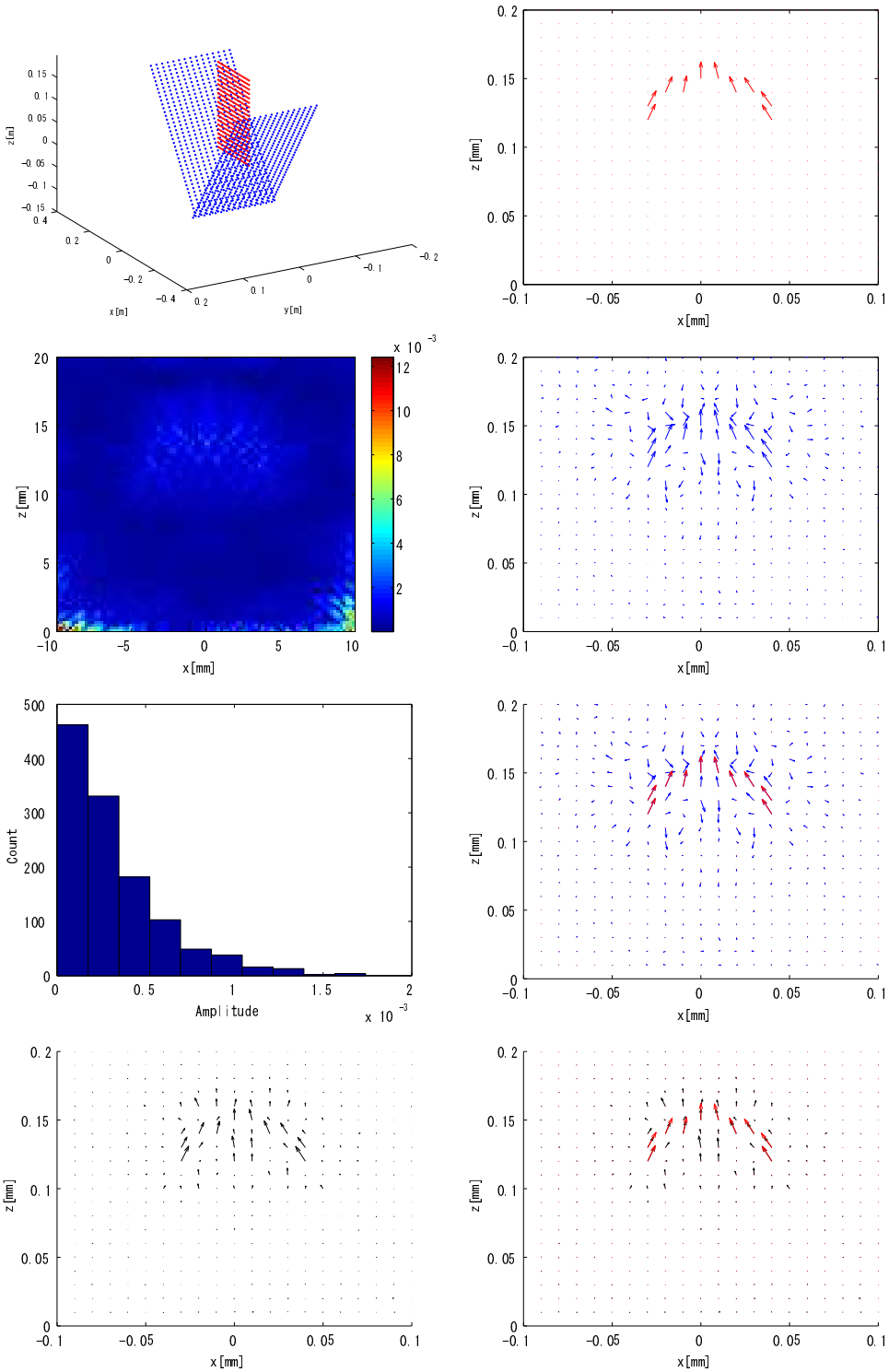

Fig. 3.14. Another example of numerical experimental results for given location condition (left top, blue dots indicate transducers and red dots indicate where the particle velocity is given) and given particle velocity distribution (right top). The obtained absolute amplitude distribution (2nd row left), imaginary part of velocity distribution (2nd row right), output amplitude histogram (3rd row left). Given and obtained velocity distribution are shown overlapped (3rd row right). Obtained acoustic intensity (bottom left), overlapped by particle velocity (bottom right). All values are normalized.

Chapter 4

Ultrasound Vibrotactile Display with a Composite Array

In this chapter, the fabrication of an actual Airborne Ultrasound Tactile Display with a composite array is described.

4.1 Hardware Construction

4.1.1 AUTD Unit Working in Arrays

As stated in the previous chapter, the array size directly limits the performance of the AUTD. Nevertheless there will be a lot of difficulties in fabricating a single large array of ultrasound transducers owing to its size and huge numbers of wirings. In addition the system with a large array lacks in flexibility of installation.

In order to overcome this problem, the multi-unit scheme was proposed and some numerical simulations was performed in the previous research [2]. As a result, it is confirmed that a wider aperture prevents the focusing in far field from being blurred. Takahashi et al. constructed a 4-unit AUTD system subsequently, which succeeded in enhancing the output power (Figure 4.1). However, due to the inter-unit gaps that played a role as spatial filter and eventually the focusing was imperfect.

We reconstructed the new multi-unit system with a newer type of array units[3]. They are fabricated so that the gap between units is 1cm. The signal transmission scheme between units has been replaced by the serial communication to reduce wiring, which originally utilized parallel cables (Figure 4.2). The designing of controlling signals inside AUTD units was originally done by Takahashi et al., which the author has modified for out new multi-unit AUTD system. The internal circuit design was a product of collaborative work by Mr. Takahashi, Prof. Shinoda, NF Corporation and the author.

The network topology of the system is depicted in Figure 4.3. Each AUTD unit has a single input signal port and three output signal ports supporting 3.3V logical differential serial signals. The incoming signal into the input port comes out from all of the output ports with a delay of 40 μs. These ports can be mutually connected via 8-pin LAN cables. Two channels of signal are available in the units, data signals and clock signals namely. They are conveyed as differential signals for the sake of stable operations, which needs two lines each. As a consequence 4 lines and the ground line is necessary for the present setup. Figure 4.3 looks like a tree topology, but the signals received by a unit go through to the next units so it is equivalent to the bus topology. Every unit receives the same signal with slight delays as the one received by the unit 1.

Every unit contains 249 transducers (T4010A1 manufactured by Nippon Ceramic Co.

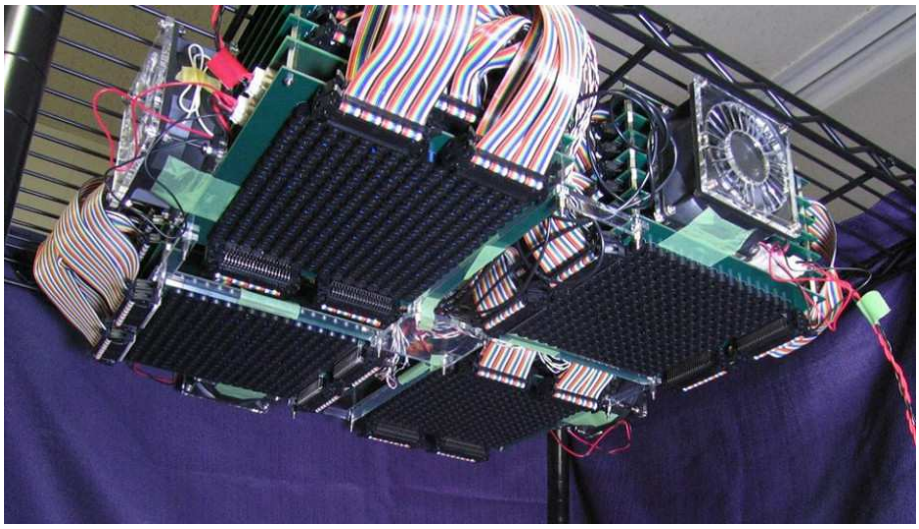

Fig. 4.1. 4 unit AUTD system fabricated by Takahashi et al. Spatial gap between arrays is seen.

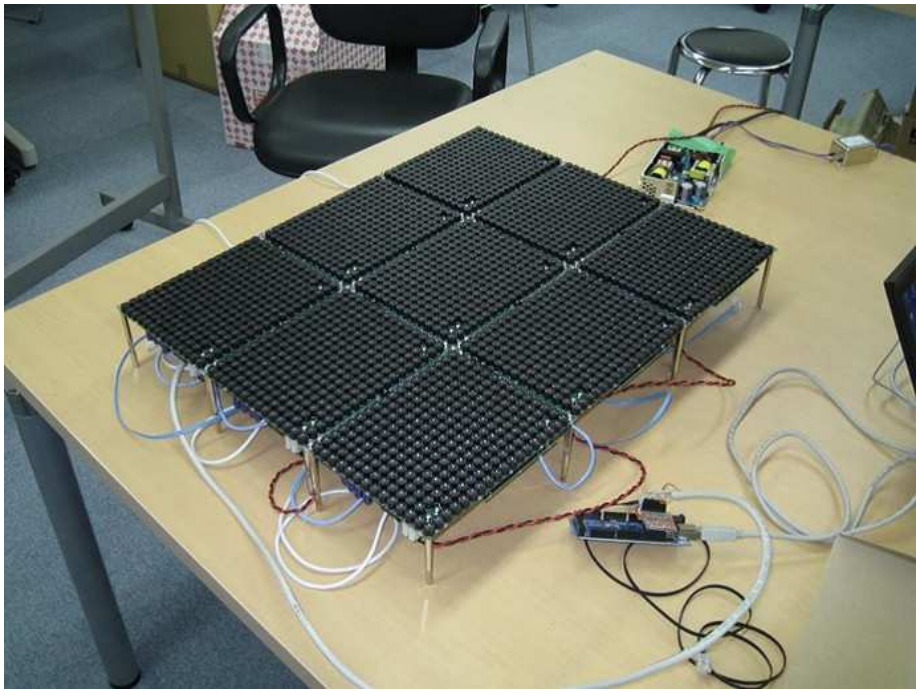

Fig. 4.2. New multi-unit AUTD system in a 3 x 3 array. The transducers are allocated in a concentrated manner.

LTD.) in a lattice of 18 transducers × 14 transducers. There are three vacant points where transducers are supposed to be for the spaces of screw heads. The array surface size is 192mm × 151.4mm. Transducers are driven by 24V peak-to-peak signals. The output SPL of a single transducer is 121.5dB at 300mm far from it in its front.

Every unit contains five FPGA(Field Programmable Gate Array)s for signal processing and output wave generation.

The one connected to its input port is what we call a 'master' FPGA of a unit. The rest four are 'slave' FPGAs, which receive the internal control signal from the master FPGA

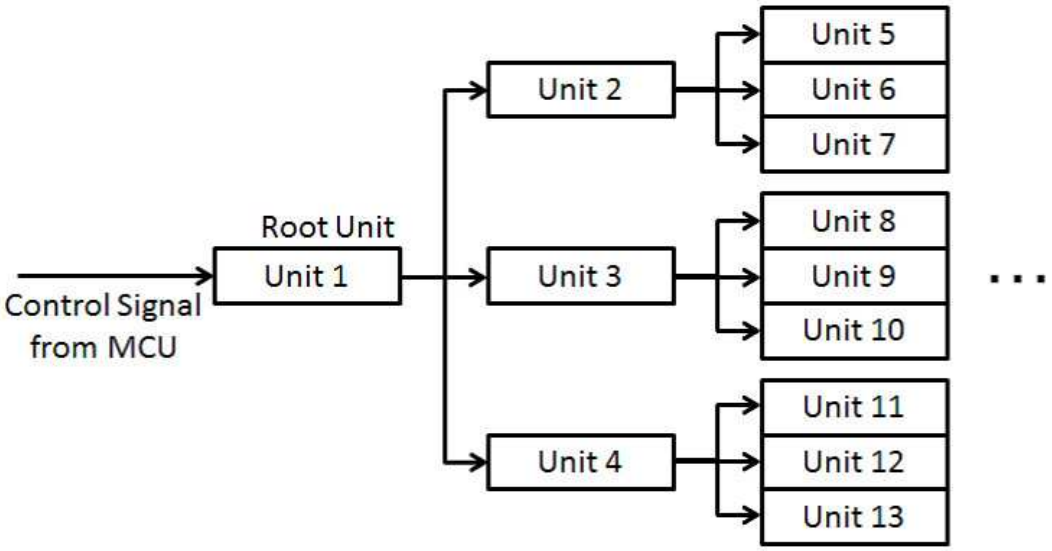

Fig. 4.3. Network topology of the multiple AUTD system

and apply driving signals on transducers of the unit (Figure 4.4).

4.1.2 Multi-unit AUTD System

To begin with, the current system produces a single ultrasound focus at the same time. The whole signal flow of the Multi-unit AUTD System is described in Figure 4.4. The system consists of a PC, a micro controller unit (MCU) and the AUTDs connected according to the network topology above.

In the following each component is described along with its allocated function in the system.

PC

The PC is supposed to send signals containing all of the necessary information for controlling the displayed focus. It transmits controlling command to the MCU via COM serial ports of the PC. Major control commands are:

- Output enable/disable: it generates or removes an ultrasound focus
- Position Set: it changes the focal position
- Send Waveform: it stores the temporal waveform of vibrotactile presented to the internal memory of the MCU
- Select Waveform Track: it switches the track of waveform stored in the MCU
- Configuration Command: it is necessary for spatial calibration, detailed description is given in the following

These commands are implemented as a software library. Since the system produces temporally vibrational intensity change in the focus by the waveform stored in the MCU, the PC has to waveform information to the MCU in serial packets.

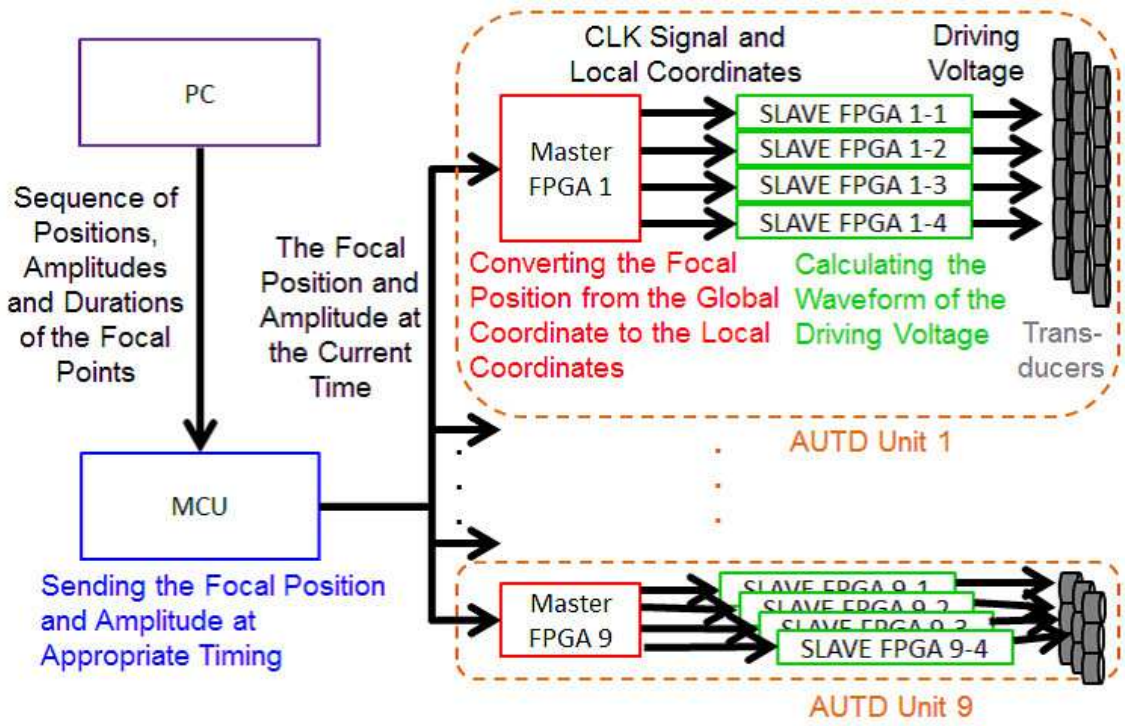

Fig. 4.4. Signal Flow of the multiple AUTD system

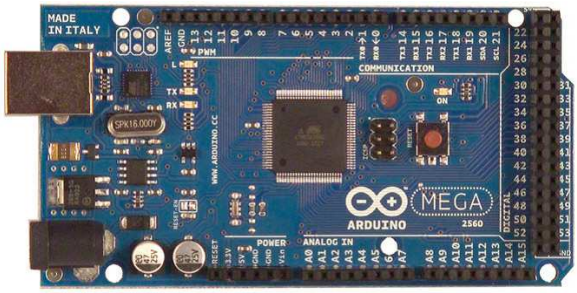

Fig. 4.5. Arduino Mega 2560 as the MCU in the system

MCU

The PC and the AUTDs are intermediated by the MCU. The MCU is what gives accurate control of temporal signal emission conveyed to the FPGAs of AUTD. In the present system Arduino Mega 2560 was applied (Figure4.5). It is endowed with 115200 bps serial communication port available for PC output, 8 KB of SRAM and 16 MHz internal clock. The current setup stores at most three different waveforms in the MCU. The MCU creates two pairs of differential serial signals at 1 Mbps emitted from the digital output ports on it. According to the received commands from PC, the internal state of MCU is changed. Internal clocks of MCU provides time counter, to which the internal state is also related. The timing of output signal is accurately controlled to the pace of 1kHz and is determined by the internal state of MCU from moment to moment. The focal positions and focal amplitude is included in this signal. The output signal voltage is changed with a conversion board before it reaches the master FPGAs in AUTDs.

Master FPGA
The master FPGA receives the serial signals. As soon as it receives signals, according to its global position and posture, the received focal position in global coordinates is converged into local coordinates of individual unit. After conversion, the focal position and intensity is conveyed to every Slave FPGA at the same time. In case that configuration signals are received, its internal memory to hold information of its position and location is renewed.

Slave FPGA
The Slave FPGAs directly control the output driving voltage on every transducer. This voltage waveform is determined by the internal clock of AUTD unit and the focal information. At most 64 transducers are under control of each Slave FPGA. The transducers are connected to 24V voltage driver ICs whose input voltage is provided by the HIGH-LOW signal of Slave FPGA. No D/A converter is currently implemented in the AUTD circuits. Instead the pulse width modulation technique is employed for output amplitude control over rectangle pulse of driving voltages. Control of duty ratio of driving pulse is described in the following section. The Slave FPGA calculates the phase delay on each transducer from its position and that of the focus.

4.1.3 Coordinate Conversion

The focal position transmitted from PC is defined in the global coordinate common to all of the units due to the network topology in which all units receive the same signals. On the other hand, the actual focusing of each unit is done based on the phase calculation of transducers using the local coordinates with in the unit. Hence the coordinate conversion from global one to local one is necessary. The position of a unit is defined as a representative point of it such as the center of the array. Its posture is for example defined as the direction of the normal vector of the array surface. The local coordinates within the unit is supposed to have three, u, v, and w axis. The direction of three axises determines the unit posture. Let $\boldsymbol{r}_l$ be the position of a unit and $\boldsymbol{e}_u, \boldsymbol{e}_v, \boldsymbol{e}_w$ be the unit vector representing the local u, v and w axis (Figure 4.6). Note that all vectors above is given in the global coordinates with x, y and z axis. With the definitions above, the global focal position $\boldsymbol{r}_f = [x_f \ y_f \ z_f]^T$ is translated into the local focal position $\boldsymbol{r}_f^{\text{Local}} = [u_f \ v_f \ w_f]^T$ by the following conversion:

$$\boldsymbol{r}_f^{\text{Local}} = \begin{bmatrix} u_f \\ v_f \\ w_f \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_u^T \\ \boldsymbol{e}_v^T \\ \boldsymbol{e}_w^T \end{bmatrix} (\boldsymbol{r}_f - \boldsymbol{r}_l). \tag{4.1}$$

4.1.4 Design of Control Signals

Signals traveling within the system are categorized into three layers, command signals from PC to MCU, focal position and intensity signals from MCU to Master FPGA and internal signals from Master FPGA to Slave FPGAs.

Command Signals from PC to MCU
Command signals from PC to MCU may contain information of focal position, local coordinate initialization, output enable command and waveform data. Each packet consists of 2 bytes and the whole data is composed of 1) start bytes, 2) header bytes denoting the data size and the data type (focal position, local coordinates, waveform or output enable

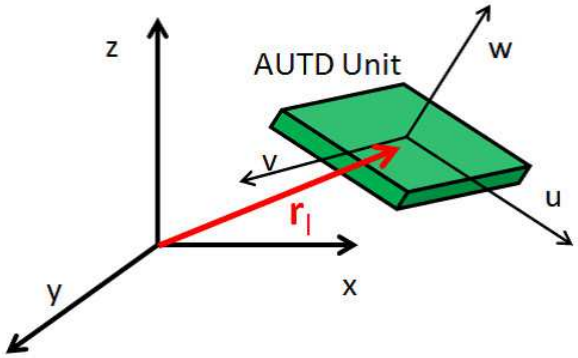

Fig. 4.6. Conversion from the global coordinates to the local one.

Fig. 4.7. Structure of packets from PC to MCU.

control), 3) data packets and 4) end bytes (Figure 4.7). The data range is confined from -32767 to 32766 so that no packet takes the value of the start bytes (0x8000) and end bytes (0x7FFF). Currently no checksum function is implemented for this communication signal. If the data transmission channel is sufficiently reliable and there is an urgent need for low latency, interactivity of the communication between PC and MCU can be omitted. The possible smallest latency with the system for creating or switching positions of a focus is about 4 ms with this packet structure (Figure4.8). This latency is divided into a data transmission and interpretation latency and sound propagation latency due to the finite sound velocity. (Figure4.9) shows the driving voltage to the transducers and generated ultrasound waveform at a focal point part from the array surface by 300 mm. Assuming the sound velocity to be 340 m/s, the propagation latency is estimated to be about 0.88 ms, which is consistent with the observation in (Fig.4.9). From these results, it is estimated that the internal latency of the system is about 3 ms.

Focal position and intensity signals from MCU to Master FPGA

What the MCU sends to the Master FPGA is only instantaneous information: the focal point x, y, z, focal amplitude p and ID, the unit ID of the current time, from moment to moment. This signal is sent to the Master FPGA at the temporal rate of 2 kHz, precisely. Each data is included into packets of 2 bytes (16 bits). At the head of the serial packets, the starting bytes are added (32 bits of HIGH level signal). Hence the total size of the packets is $32 + 5 \times 16 = 112$bits (Figure 4.10). If one bit takes less than 4 $micro$s, packet transmission at 2kHz can be maintained. The value of focal smplitude p is determined frame by frame according to the time counter inside the MCU and the stored waveforms. There should be explanation on how to deal with the local coordinate initialization signals.

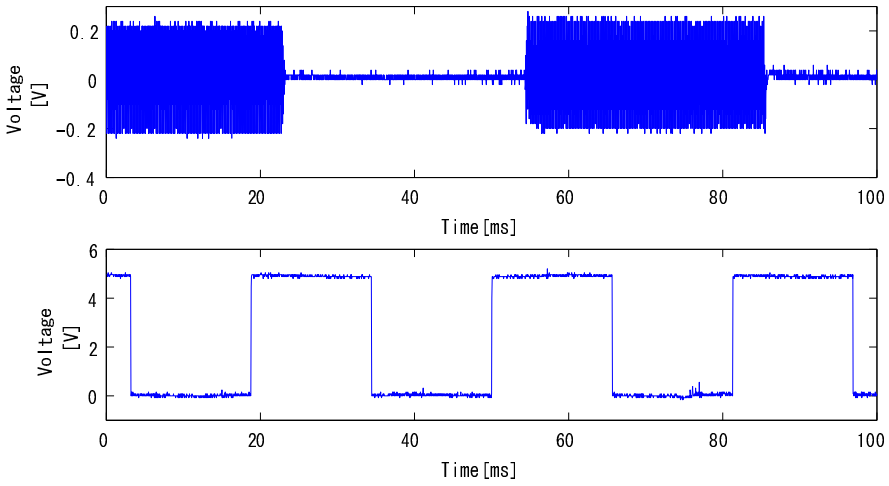

Fig. 4.8. Sound pressure at the focal point (Upper) and trigger voltage of command transmission onset switching two focal positions alternatively (Lower). Approximately 4ms of latency between the rising edges of trigger voltage and the onset of recorded sound waveform is observed.

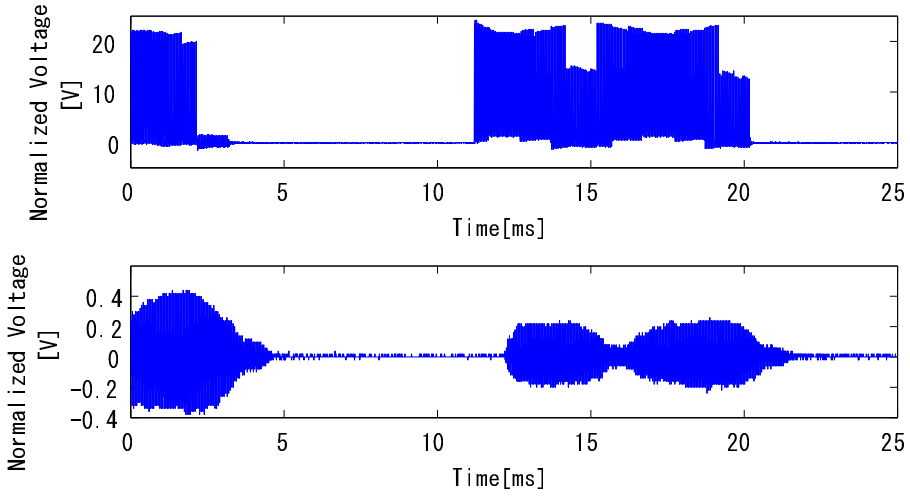

Fig. 4.9. Flickering driving voltage applied to one of the transducer and sound pressure at the focal point (Lower). Approximately 1ms of latency between the rising edges of driving voltage and the onset of recorded sound waveform is observed.

In case the MCU receives that command, output value of ID, the last two packets, is sent to Master FPGAs of all units. For other command signals, the value of ID stays zero. As described in the following, each Master FPGA holds its own unit ID and when the received value of ID in the signal is coincident with its unique ID, it interpret the received signal as the local coordinates initialization information for it. For this case, the value of p includes the additional information for initialization: which one the other data (x, y, z) denotes among $\boldsymbol{r}_l, \boldsymbol{e}_u, \boldsymbol{e}_v$ and $\boldsymbol{e}_w$.

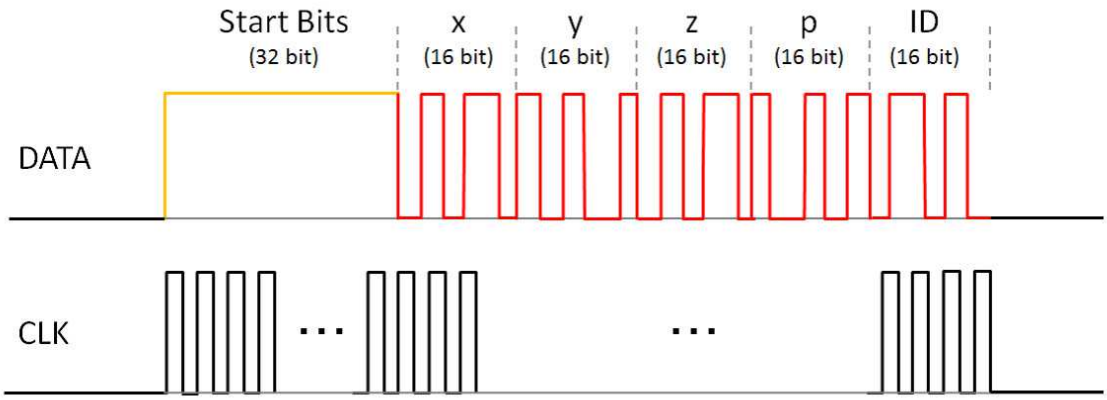

Fig. 4.10. Structure of serial signals sent from MCU to Master FPGA.

Internal signals from Master FPGA to Slave FPGAs
Internal signals within an AUTD unit are conveyed via 16-channel parallel wiring. The signal contains focal position and amplitude converted to its own local coordinates. This conversion is instantaneously done by the wiring inside the Master FPGA. As stated above, no internal signal is sent when what the Master FPGA receives is the initialization signal. Four Slave FPGAs receive this signal from Master FPGA simultaneously.

4.1.5 Phase Shift Calculation

The output signal of each transducer is determined by the amplitude p and its phase shift. The focal position (x, y, z) is expressed with a unit determined by $\lambda/16$ where λ is the wavelength of transducer, namely 40kHz in our setup. This is for the sake of simplified calculation of the phase delay by omitting the conversion of physical dimension of the data from the length[m] to phase[-]. Eventually the phase shift is quantized into 16 levels. One level of phase shift is equivalent to the 0.5 mm error in the sound field. The internal clock of AUTD is 25.6 MHz and it takes about 1,200 clocks for the phase shift calculation of 64 transducers, resulting in less than 50μs latency from the arrival of the internal signal. In this calculating, an iterative method for square root computing is employed. Since the iteration count is fixed, every Slave unit finishes the calculation at the same time.

4.1.6 Focusing Performance Test

We operated experiments in order to verify that the spatial profile of the radiation pressure was generated in control. We constructed 9-unit AUTD system and examined its focusing performance. The number of transducers was 2,241 in all. The array size was 576mm $\times$ 454.2mm.

We measured the spatial distribution of ultrasound amplitude in the experiment. Figure 4.11 shows the experimental setup. The 9-unit AUTD was mounted on an aluminum cabinet so that the transducer surface faced the ground. The surface was parallel to the ground. A standard microphone (Bruel & Kjar Type 4138) with a pre-amplifier (Bruel& Kjar Type 2670) was mounted on the 3D stage, whose xy-plane is parallel to the AUTD. The xy coordinates corresponded with the lattice of transducers and the z axis was vertical to the AUTD surface.

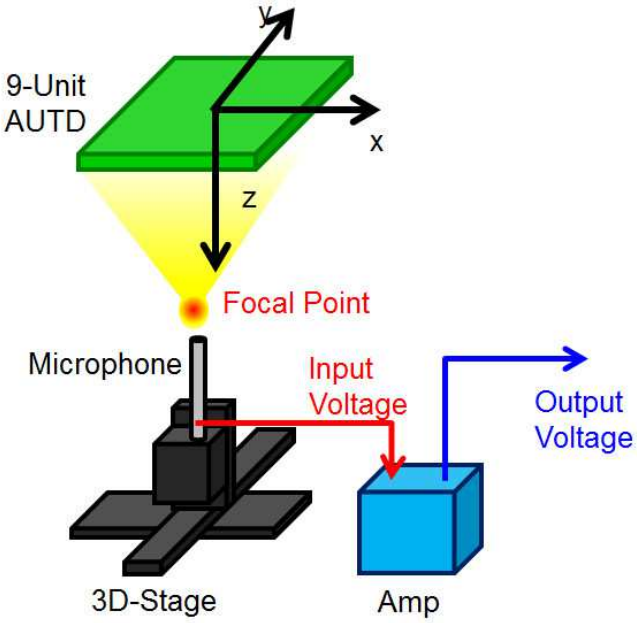

Fig. 4.11. Schematic image of the experimental setup.

The recorded voltage was amplified with a power amplifier (Bruel & Kjar Type 5935). We used the reference sound source which could provide 94dB sine wave of 1 kHz for the mapping of the recorded voltage to sound pressure. The frequency characteristic of the whole recording system was almost flat from 0 to 40 kHz. The focal point was set to (0, 0, 600mm). We measured sound pressure near the focal point and estimated the produced radiation pressure from 3.17. The parameters were set as c=340m/s, ρ_0=1.18kg/m^3 and α=2.

Figure 4.12 depicts the calculated radiation pressure distribution in 2d-plane at z = 600mm. High radiation pressure can be seen localized within the diameter of 10 15 mm. Also, the directivity of the pressure distribution can be seen since the aperture of the AUTD was rectangular. A more macroscopic distribution is depicted in4.13. This is a 1D distribution across the focal point where y and z were set to 0 and 600, respectively. The radiation pressure can be seen drastically attenuate outside of the focal regions. As a result, the output radiation pressure at the focal point was expected to be approximately 400 Pa from the measurement. The gross force measured around the focal point reached 7.4 gf.

This result of experiment shows that by creating a large aperture, proper focusing in ranges comparable to the array size is guaranteed.

4.2 Temporal Control of Sound Amplitude

4.2.1 Creating Vibrotactile Sensation with Pulse Width Modulation

The driving voltage signals of ultrasound transducers in our current setup are digital pulse waves. The transducers have a very sharp band-pass frequency characteristic of input voltage. As a result, the 40 kHz components in the input pulse voltage are derived as output ultrasound waveform. The amplitude of this 40 kHz component can be controlled by tuning the duty cycle of input pulse voltage. Suppose the frequency of the input pulse

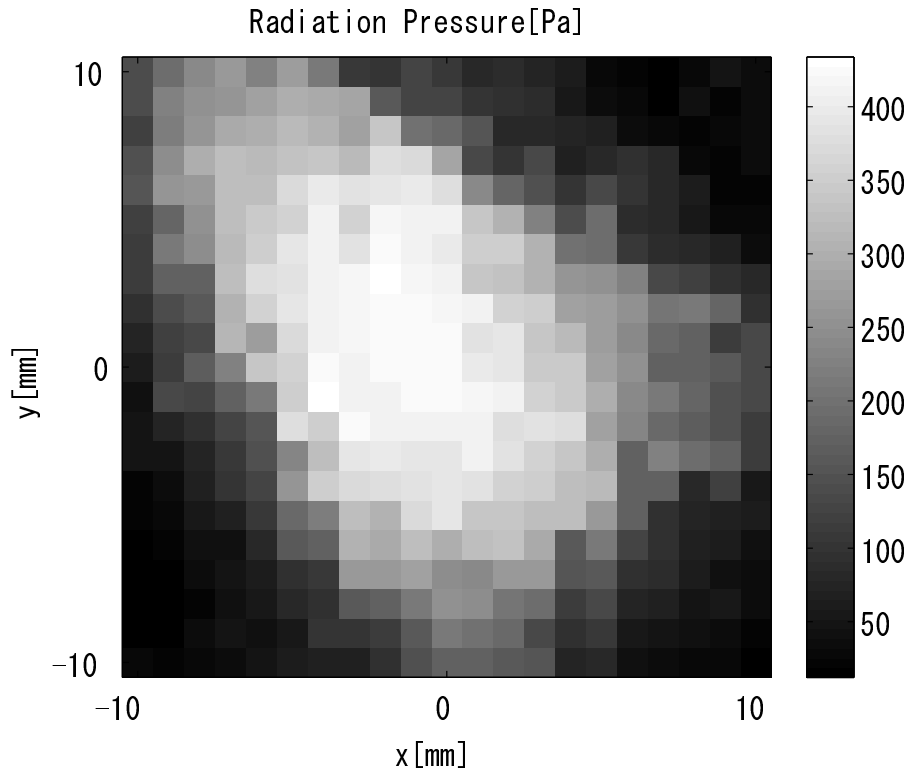

Fig. 4.12. Spatial distribution of radiation pressure estimated from amplitude measurements. The plane is parallel to the array at $z = 600$mm.

is set to 40 kHz and its duty cycle is set to d $(0 \leq d \leq 1/2)$. The RMS of output ultrasound amplitude p becomes

$$p = p_M |\sin(\pi d)|. \tag{4.2}$$

where p_M denotes the amplitude when d is equal to 1/2 [4]. Therefore, resulting waveforms can be controlled by changing the duty cycle of input voltage temporally (4.14). Once the duty cycle and time delay τ[s] determined from the phase shift of each transducer is determined, driving signal $V(t)$[V] of the transducer is given by

$$V(t) = \begin{cases} V_M, & (nT - \tau \leq t \leq nT + dT - \tau) \\ 0, & (nT + dT - \tau \leq t \leq (n+1)T - \tau) \end{cases}, \tag{4.3}$$

where $T = 1/f$[s] is the period oh the ultrasound. Here V_M holds the relationship that 40 kHz component of $V(t)$ is equal to p_M when $d = 1/2$. From 4.2 and 3.17, when the duty cycle is expressed as a time-variant function $d(t)$, the resulting radiation pressure can be expressed as

$$\begin{aligned} P_R(t) &= \frac{\alpha}{\rho c^2}\{p(t)\}^2 \\ &= \frac{\alpha}{\rho c^2} p_M^2 \sin^2(\pi d(t)) \\ &= \frac{\alpha}{\rho c^2} p_M^2 \left\{ \frac{1 - \cos(2\pi d(t))}{2} \right\}. \end{aligned} \tag{4.4}$$

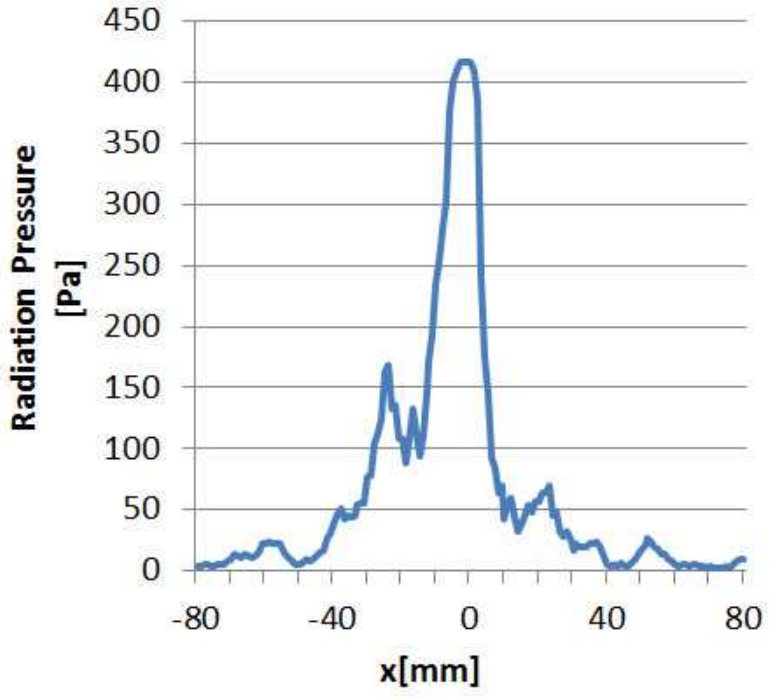

Fig. 4.13. Macroscopic radiation prssure distribution on the x-axis at $y = 0, z = 600$mm.

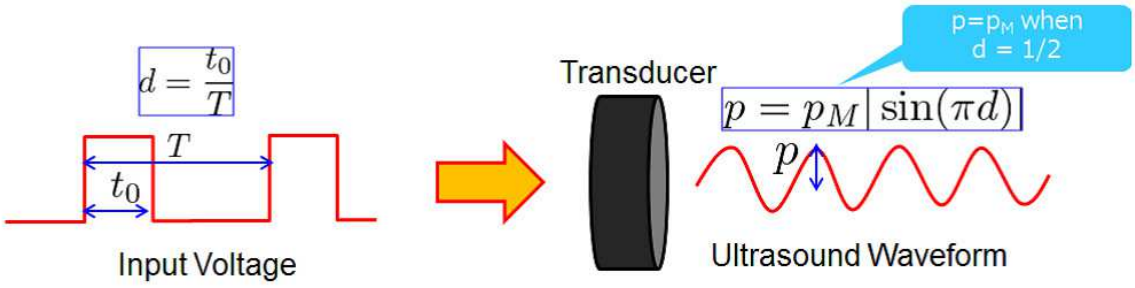

Fig. 4.14. Output ultrasound amplitude control by duty cycle variation.

Using this relation, the driving pattern $d(t)$ for a given radiation pressure waveform $P_R(t)$ can be obtained as

$$d(t) = \frac{1}{2\pi}\cos^{-1}\left\{1 - \frac{2\rho_0 c^2}{\alpha p_M}P_R(t)\right\}. \tag{4.5}$$

Since the value of cosine should be between -1 and 1, the term $\frac{2\rho_0 c^2}{\alpha p_M}P_R(t)$ must not exceed 1. This term can be considered as a normalized output waveform of radiation pressure. The relation between these three function $V(t)$, $p(t)$ and $P_R(t)$ is depicted in Figure 4.15. Thus the time-variant radiation pressure with a desired temporal profile is generated by varying the duty cycle of driving voltages in accordance with the equation above. Due to the low-pass characteristics of human skin and mechanoreceptors, the squared envelope $P_R(t)$ is what we feel as vibrotactile sensations[3]. Note that with this principle negative pressure cannot be generated. For the improvement of vibrational quality in some case,

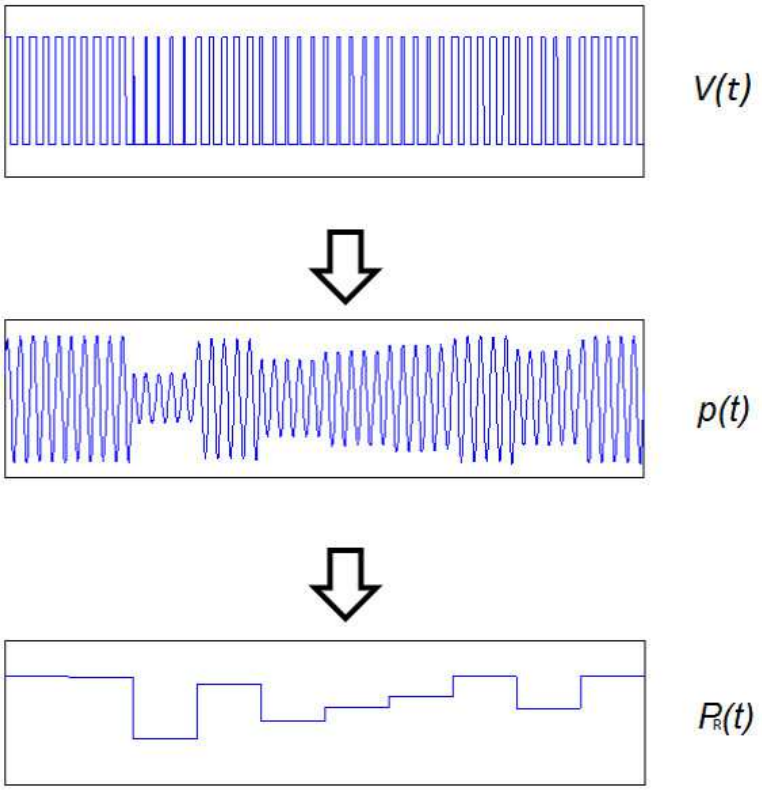

Fig. 4.15. Relation between driving signal $V(t)$, sound amplitude $p(t)$ and time-variant acoustic radiation pressure $R_P(t)$.

adding offset radiation pressure so that the center of the waveform becomes non-zero might be necessary. In other cases presenting feel of active touches on the vibrating object, rectification of the object vibration waveform might be sufficient for the reproduction of the phenomenon. As referred to above, under some conditions the suction pressure has been demonstrated to indistinguishable from positive pressures[46]. How to design the optimal vibration to tactually convey the desired physical situation at the highest fidelity remains as an essential future task in the present research.

4.2.2 Improvement in Duty Cycle Quantization

Since the output power is much improved as mentioned, it has become more realistic to create the variety of displayed vibrotactile sensation by tuning the temporal driving pattern appropriately. In particular, ultrasound envelope of a high sampling rate and a fine quantization is needed. In order to realize this, the driving pulse $d(t)$ has to vary quickly and its dynamic range has to be as wide as possible. Our current AUTD units include 25.6 MHz internal clocks and the driving pulse signals are generated from them. Thus the duty cycle of 40 kHz pulse $d(t)$ can be quantized up to 640 levels. From (4.2) and (4.4), it is shown that the radiation pressure $P(t)$ is quantized up to 320 levels, which has been improved from 8 levels in the previous prototype[4]. As stated in the previous section, the current system can switch the duty cycle fast enough to generate vibrotactile sensation with the temporal resolution of 0.5 ms (2kHz) with the MCU.

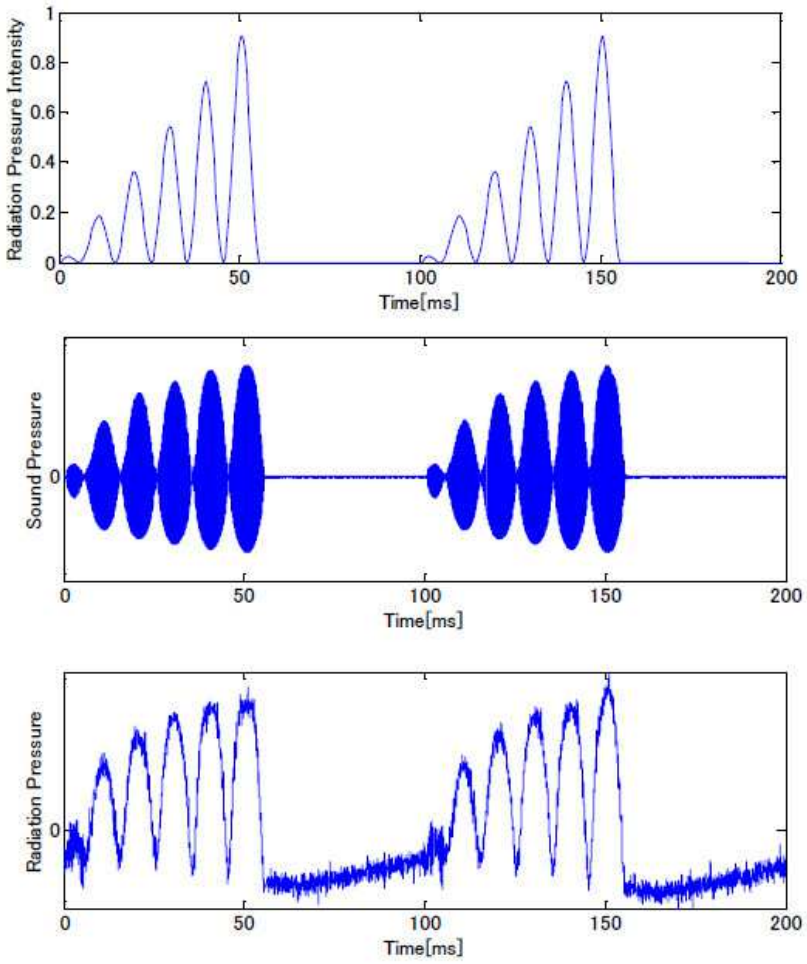

Fig. 4.16. An example of input signal of radiation pressure (upper), generated sound pressure (middle) and radiation pressure (lower)[3].

In the previous prototype the waveform was determined by the 16bit-decoder describing the output waveform for every possible pattern of 16bit phase shifts and 8bit amplitudes. Instead of this implement, the digital output voltage is determined in the present system according to (4.3) from the given amplitude and calculated phase shift. An example of recorded sound pressures and radiation pressures generated from an input waveform pattern is shown in 4.16. An Asymmetric sound pressure waveform, which caused radiation pressure, is seen in the figure 4.16. This waveform distortion is a symbolic phenomenon of nonlinear acoustics caused by high sound intensity. The radiation pressure amplitude depicted here is derived from the measured sound pressure amplitude by eliminating its 40 kHz or higher components. In the figure the generated radiation pressure has its temporal profile similar to the input signal, yet not completely the same. There are several reasons possible. First, what is extracted as radiation pressure from the recorded sound pressure might not be the exact radiation pressure exerted to the microphones. This possibility suggests that there is a room for measurement improvement. The next cause might be the distortion of waveform which is more complicated than that speculated by (3.17). More detailed physical modeling of radiation pressure can help verifying this hypothesis. In addition there may be still other nonlinear acoustic phenomena affecting this measurement such as the acoustic streaming. There was a fluctuation of measured amplitude in much

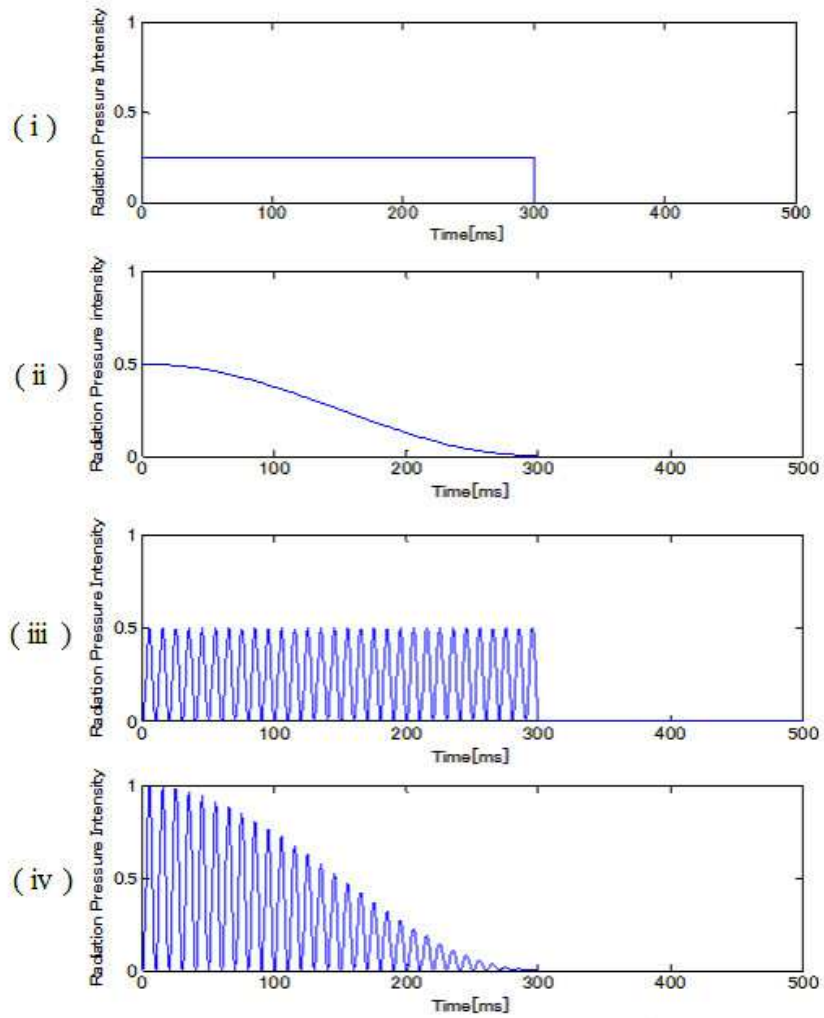

Fig. 4.17. The four vibrotactile stimulation patterns added on subjects' skins.

a longer time scale compared to the ultrasound period (up to several seconds), which indicates the possibility of other phenomenon existing.

4.2.3 Vibrotactile Sensation Identification Test

In this experiment we generated four different vibrotactile patterns and human subjects distinguished them only by touch. Figure 4.17 shows the temporal profiles of the four vibrotactile patterns displayed on subjects' skins. The time integration of applied force generated from each pattern was adjusted to be equal so that the subjects would not be able to identify the displayed patterns by their average force intensity. A 100-Hz sine wave is modulated by each envelope in the pattern (iii: nonattenuating vibration) and (iv: attenuating vibration). The envelopes of the patterns (i: static pressure with impulsive rising and falling edges, like a mouse click) and (iii) are identical and the same goes for the patterns (ii: static pressure starting with rising edge and attenuating end) and (iv). We chose these patterns in order to assess how accurately subjects can distinguish 1) static and vibrating stimuli with the same envelope, and 2) envelopes of stimulation waveforms when the carrier waveforms are the same. The durations of all patterns were

Fig. 4.18. Picture of the identification experiment. On the top of the cabinet was the 9-unit AUTD mounted.

set to 300 ms, which we considered a typical duration of a mouse click. Six subjects were asked to identify the vibrotactile stimulation which the AUTD generated on their palms among the four patterns described above. The stimulation was in a non-moving spot set to a time-invariant identical position. The distance between the subjects' palms and the AUTD was approximately 680 mm. The waveforms were printed on a piece of paper and the subjects were able to see it throughout the experiment. Before the identification task, the subjects received each stimulus pattern on their palm being informed of the stimulus number. Since the AUTD made characteristic noises when generating focused ultrasound, the subjects' wore headphones and heard white noises to mask those sounds during the task (Figure 4.18). Throughout the experiment, no visual clue such as the obvious deformation of the skin surface was seen. The four patterns were displayed 10 times each, 40 times in all. The patterns were displayed repeatedly until the subject gave an answer. The sequence of the patterns was randomized. All subjects were male and their ages were between 22 and 47.

Figure 4.19 shows a graph of the correct answer ratios of the experiment. The error bars here indicates standard deviations of the correct answer ratios. The pattern (i) and (ii) were identified much more accurately than the rest of them. Some of the subjects claimed that the perception of the falling edges was a clue in distinguishing the pattern (i) from (ii). One of the other subject argued that he could feel the attenuation of the pressure in the pattern (ii), which became his clue for differentiation of the pattern (i) and (ii). Figure 4.20 shows a confusion matrix between the subjects' answers and actual displayed patterns. The number in the table indicated the numbers of times of answers. The table indicates that all subjects were able to differentiate the pattern (i) and (ii) from (iii) and (iv). It can be said that the differentiation between the pattern (i) and (ii)

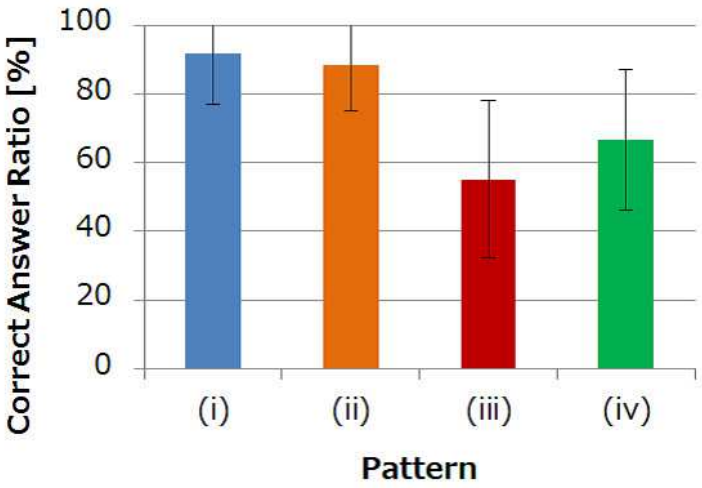

Fig. 4.19. The correct answer ratio.

Answered / Displayed	(i)	(ii)	(iii)	(iv)
(i)	55	5	0	0
(ii)	7	53	0	0
(iii)	0	0	33	27
(iv)	0	0	20	40

Fig. 4.20. Confusion matrix between the subjects' answers and actually displayed patterns.

were easier than that between (iii) and (iv). Distinction between the falling edges of the envelope of (iii) and the attenuation pattern of (iv) felt vague.

4.3 Generating Vibrotactile Stimuli from Recorded Signals

4.3.1 Converting Recorded Signals into Vibrotactile Waveform

In the experiment above, the radiation pressure patterns were created artificially through a mathematic manner. As a result, the displayed stimuli are something convenient for analysis of the experimental result, but not the ones with attractive of natural tactile texture. As demonstrated in examples in Chapter 2, vibrotactile stimuli created through proper procedure can achieve the reproduction of texture with high fidelity. In Chapter 2, two major strategies to produce such a vibrotactile sensation is refered to. One is reproducing recorded real vibration and another is utilizing a well-constructed stochastic

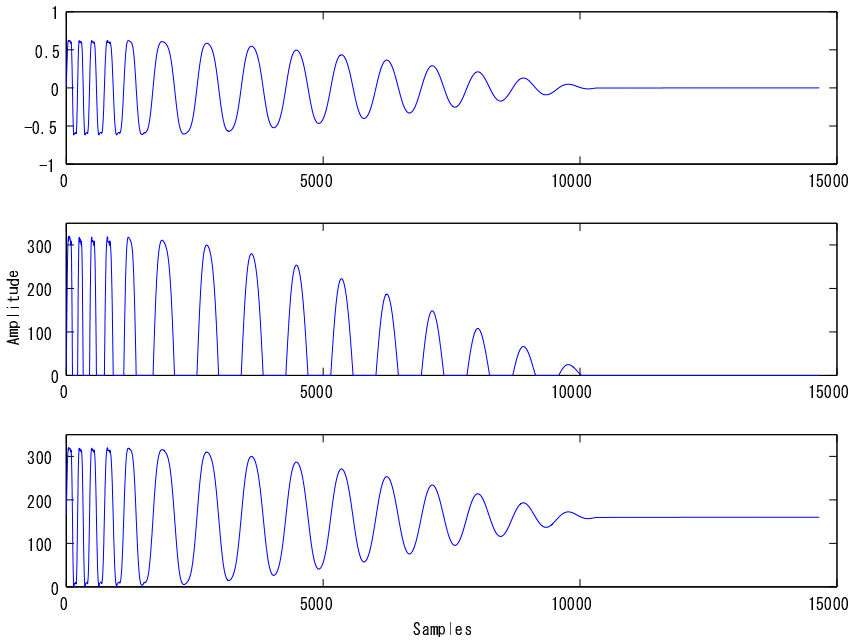

Fig. 4.21. Waveform extraction from audio signal. The original signal (Upper), extracted signal with half-rectification (Middle) and that added with DC offset (Bottom).

model. However the latter method is still at a developing stage since it can only be applicable for single-point contacts through held apparatus. Thus the author describes the result on creating vibrotactile stimuli through the former method.

The radiation pressure $P_R(t)$ can be generated from a recorded signal. by resampling and rectifying the negative part of the signal or adding some offset to it for expression of its negative part (4.21). After $P_R(t)$ is determined, the duty cycle sequence $d(t)$ is obtained by (4.4) as above.

The author has tested the tactile feelings generated from a set of time variant acoustic radiation pressure extracted from audio signals.

As an overall tendency, extracted waveform through half-wave rectifications displayed intense tactile feeling to users. On the other hand, the fineness of displayed vibration seems to be better in those with DC offsets. Figure 4.22 shows examples of generated vibrotactile waveform in duty cycle domains. The top left figure corresponds to a waveform with impulsive feelings. The attenuation of almost a single-frequency sinusoidal wave is seen in it. Thanks to that, a feeling of being hit something solid and lite in eight with small reverberate reactions. This waveform is extracted from an artificial bongo waveform. Second one in the top right figure is obtained by rectifying another bongo waveform in which distortion of waveform can be seen. This waveform gives the feeling of touching something elastic is vibrating on the skin surface. When generated to the fingertip, the author feels as if he touched a vibrating rubber ball. The third one depicted in the bottom left figure is actually not extracted from audio signals. A rectangle pulse is seen followed by sinusoidal wave in the figure. This is designed so that it resembles a physical behavior of a buzzer button which gives impulsive clicking reaction as they are pushed followed by buzzing vibration inside it. The last one shown in the bottom right figure is an exciting one. It is obtained by rectifying an audio signal of the bass drum. Something rapturing is vividly felt on the skin surface with this waveform. Among the four vibrational waveforms,

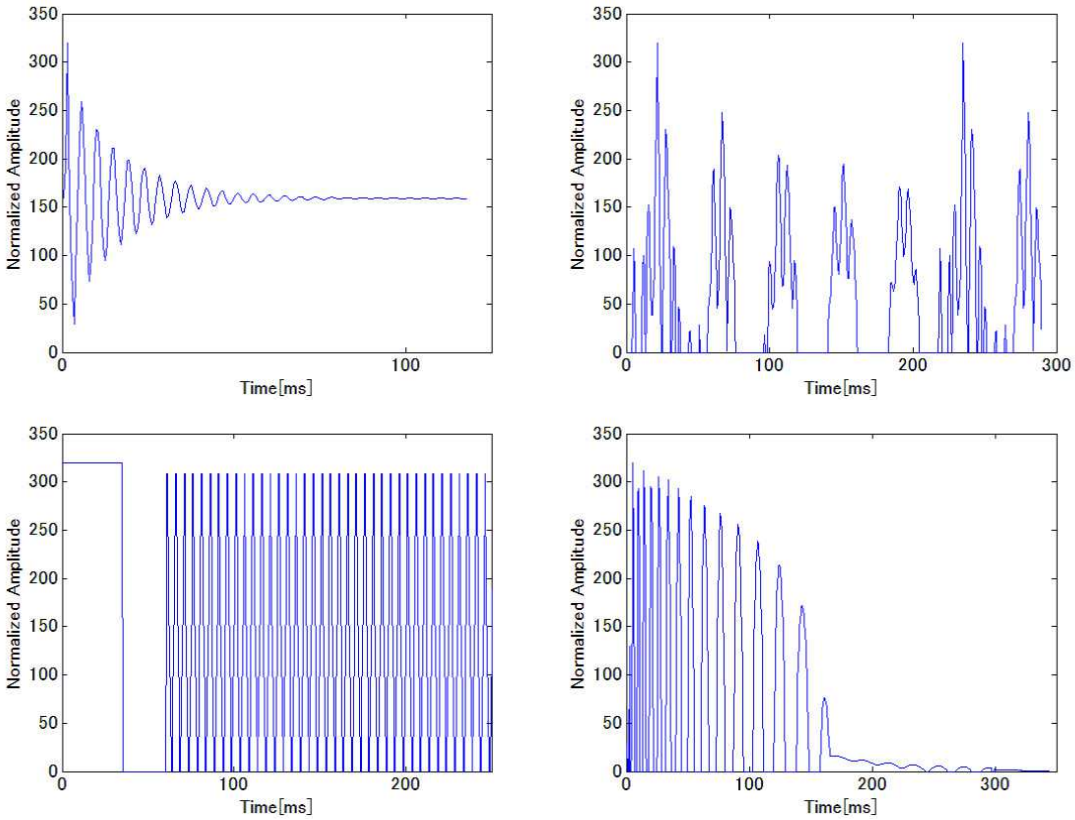

Fig. 4.22. Examples of designed vibrotactile waveforms. Impulsive(Top Left), Trembling(Top Right), Buzzing(Bottom Left) and Rapturous(Bottom Right).

the last one provides the most vivid tactile sensations due to its frequency characteristics. These waveforms are currently used in the 'Tactile Projector' demonstration in Chapter 6[63].

Chapter 5

Reduction of Accompanying Audible Sound

The current AUTD has another problem in real world operations. It is the generation of audible sound accompanying with creation of radiation pressure focus. The existence of the audible noise limits the range of applications. It is obvious that it cannot be used in situation which needs silence. Also, personality of invisible tactile stimuli conveyance is no available since it is heard. Although there have been several hypothesis for the cause of the noises, this problem has not been discussed in greater detail. In the chapter it is presented that changes in spectrum structure by the ultrasound amplitude modulation is one of the key mechanisms to generate audible noises. The author proposes a reduction based on ultrasound envelope smoothing by gradually shifting the duty cycle of driving voltages. In the last part of the chapter, experimental verification is presented.

The issues described in this chapter were investigated in cooperation with Mr. Moto Yoshioka, a 4th-grade student in the faculty of engineering in University of Tokyo at the time of the thesis submission, under the supervision of the author.

5.1 Generation of Audible Sound

As presented in the Chapter 4, the amplitude modulation is employed for generating various vibrotactile sensations. However, this modulation is one of the critical causes to generate audible noises. Let $p_c(t)$ the carrier waveform emitter from transducers. For simplification, it can be assumed that $p_c(t) = \cos(\omega_c t)$, $\omega_c = 40{\cdot}2\pi$ rad/s. In practice $p_c(t)$ is multiplied by its amplitude $< p >$, which holds (3.17). It should be noted that (3.17) does not take the temporal component of amplitude into account since it takes time average of amplitude for extracting the envelope $< p >$. The time average $< p >$ is considered to temporally vary in a larger time scale than the period of the carrier waveform in generation of vibrotactile stimuli. Since the time average envelope $< p >$ is represented as a temporal function $p_e(t)$. Hence the amplitude can be rewritten as

$$p(t) =< p > p_c(t) = p_e(t)p_c(t). \tag{5.1}$$

The behavior of $p(t)$ is well understood in the frequency domain. Let $F(\omega), F_e(\omega)$ and $F_c(\omega)$ be the Fourier transforms of $p(t), p_e(t)$ and $p_c(t)$, respectively. Then, the following well known relation is derived:

$$F(\omega) = F_e(\omega) * F_c(\omega), \tag{5.2}$$

where $*$ denotes the convolution operator.

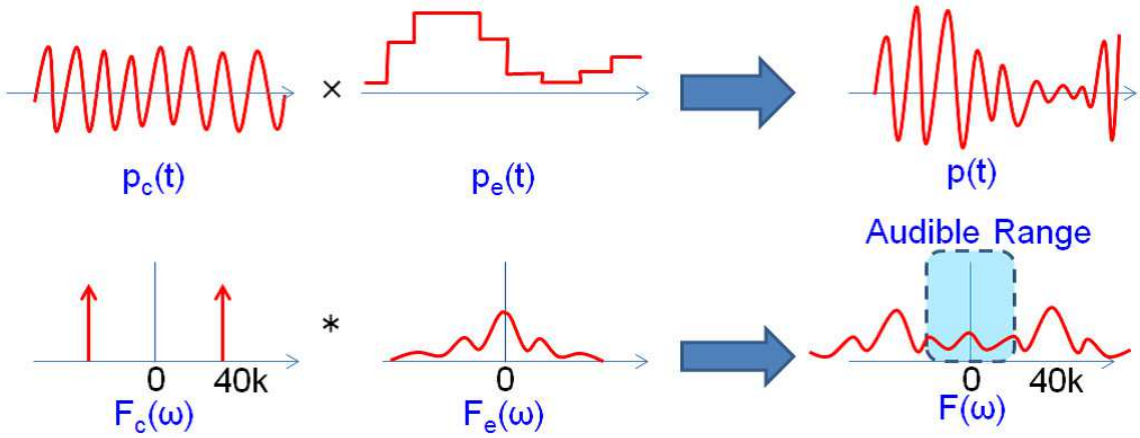

Fig. 5.1. Frequency shift by amplitude modulation.

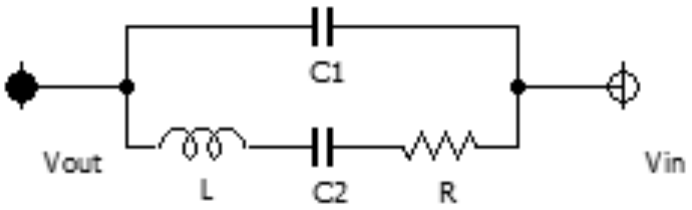

Fig. 5.2. Equivalent circuit of ultrasound transducer T4010A1.

Here, what is expected from the equation above? Since $p_c(t) = \cos(\omega_c t)$, 5.2 can be rewritten as

$$F(\omega) = \sqrt{\frac{\pi}{2}} \{F_e(\omega + \omega_c) + F_e(\omega - \omega_c)\}, \tag{5.3}$$

This is understood that frequency shift of the envelope spectrum by ω_c is caused (Figure 5.1). There is no problem if the envelope $p_e(t)$ is a band-limited signal. However, the actual envelope of the system described in Chapter 4 is considered to have spiky edges as depicted in Figure 5.1. Signals with spiky temporal profiles in general are regarded as wide-band signals containing high frequencies. Hence the resultant spectrum of ultrasound seeps to the audible ranges.

In Figure 5.2, the equivalent circuit of T4010A1, transducers used in the system is depicted. The circuit parameters are, $R = 700\Omega$, $L = 80\text{mH}$, $C_1 = 2700[\text{pF}]$ and $C_2 = 200[\text{pF}]$. For the analysis of transient response, transfer function of the circuit

$$G(s) = \frac{Rs}{Ls^2 + Rs + \frac{1}{C_2}} \tag{5.4}$$

is simulated by a second-order Chebyshev IIR filter given as:

$$G'(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2^{-2}}, \tag{5.5}$$

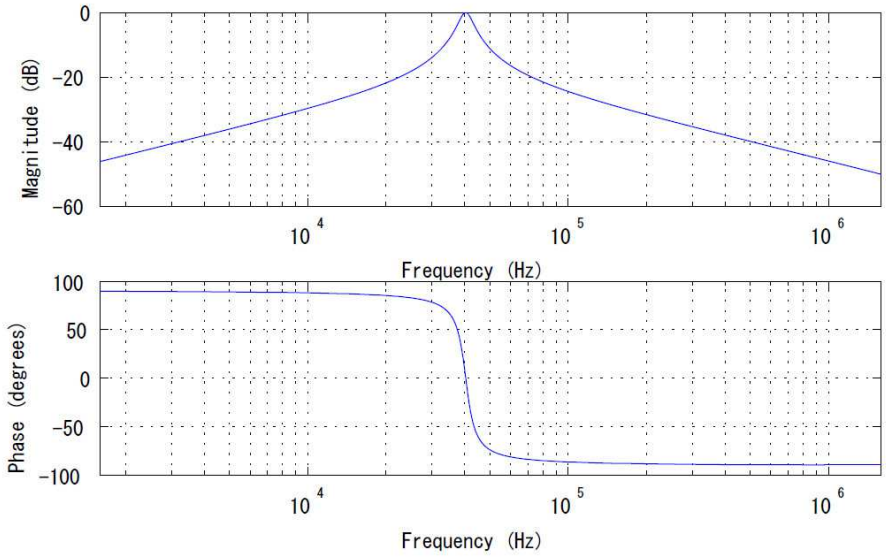

Fig. 5.3. Bode plot of a digital filter to simulate the behavior of T4010A1.

where $a_1 = -1.9987, a_2 = 0.9988, b_0 = 0.0006147, b_1 = 0, b_2 = 0.0006147$. The sampling frequency is set to 25.6 MHz. In Figure 5.3, the frequency response of the simulative filter is depicted. By applying this filter for input driving signals, their simulative output waveforms are obtained.

Figure 5.4 shows the output signal for input signal of 40 kHz rectangle pulse series with its duty cycle changed every constant samples. As a result the steep changes in the envelope can be seen. To clarify the effect of these changes on audible sound generation, the down-sampled signal is shown in Figure 5.5. The resampling frequency is 22,050 Hz. The spiky pulse is clearly seen at the discontinuous changing point of the envelope though much weakened in power compared to that of the original signal. The spectrogram of this resampled signal shown in Figure 5.6 indicates that at the moments of jumpy envelope variations, noises with broad frequency bands are generated. It is true the power of these noises might be negligible in a single transducers, but the current system employs a hundreds of transducers at least. It is natural if the power of this relatively small component in the output waveform achieves an audible level. Thus it is almost certain that the discontinuous change in the wave envelope causes audible noises.

5.2 Smoothing the Ultrasound Envelope with Discrete Sampling Filtering of Duty Cycles

The most straight forward way of solving the issue above is to add a high-pass filter to cut off the audible components from the output waveform. However the circuit of AUTD is so complicated that adding that sort of analog filter is troublesome. For that reason in this chapter an approach to smooth the envelope is taken.

Figure 5.7 indicates a simulative result of generating output waveform by two different duty cycle series. A smoother envelope is seen in a signal generated from a smoother sequence of duty cycles generating driving pulses. In this case duty cycles vary for every pulse.

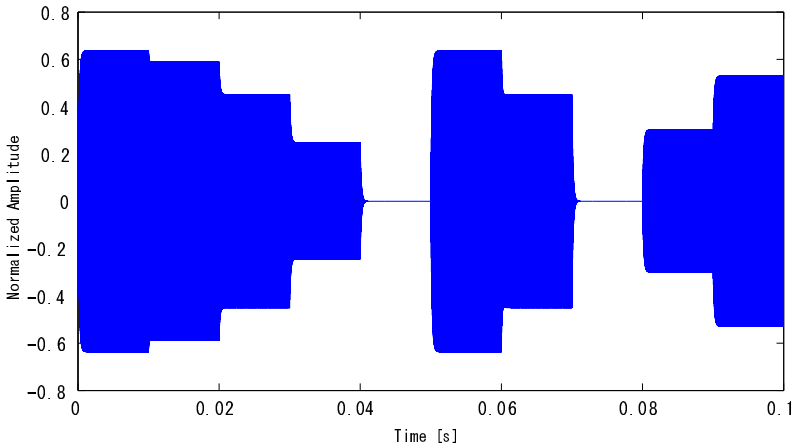

Fig. 5.4. Output waveform for a series of rectangle pulse with discontinuous duty cycle changes.

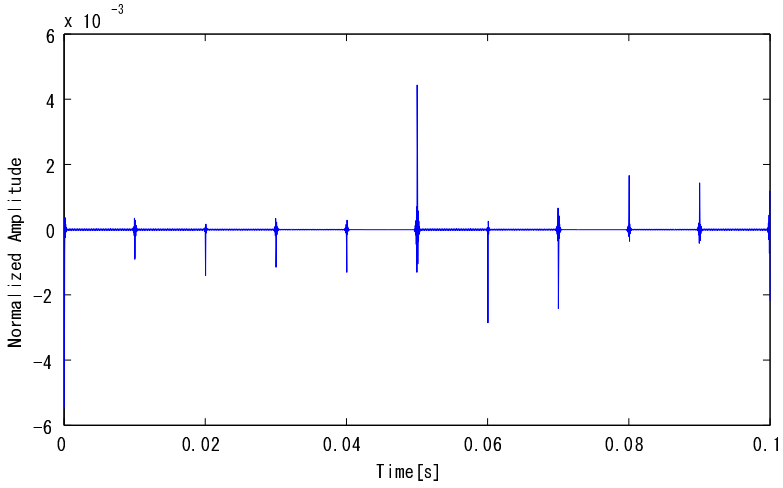

Fig. 5.5. Down-sampled signal of that in (Figure 5.4).

The current system guarantees 2 kHz vibrotactile stimuli. It means that the duty cycles vary every 0.5 ms. By the way, the ultrasound of 40kHz is generated a train of rectangular pulses whose frequency is 40kHz. This fact means that the duty cycle of a pulse train inside the FPGA changes at every 20 pulses. This is considered to cause discontinuity in the envelope. Of course, to be precise, this discontinuity cannot be completely vanished. Nevertheless, if the duty cycle changes at every pulse, it is expected that the envelope becomes smoother and the audible sounds become weaker.

The MCU in the system has a 16MHz internal clock, which cannot afford to send signals of higher frequency. Instead, the Master FPGA received the focal intensity as a form of duty cycle, is implemented with an internal digital filter affecting on the duty cycle of the driving pulses. In other words, an interpolation function of temporally sparse duty cycle value sent by MCU is implemented by the digital filter.

As the filter, an FIR filter is implemented. This is because FIR filters guarantee stable output since there is no feedback of output signals. This property is important in the system because numerical errors such as rounding errors might be accumulated during the operation, which would cause unwanted outputs. Since the system is supposed to

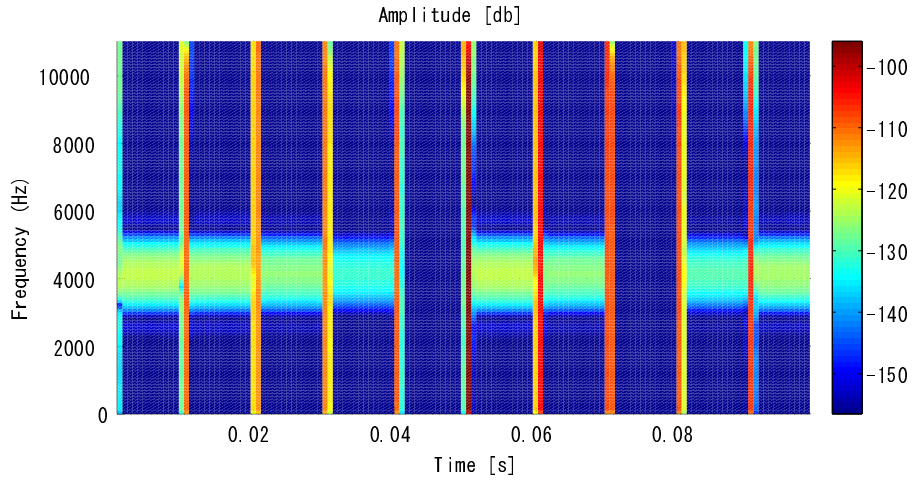

Fig. 5.6. Spectrogram of a signal depicted in (Figure 5.5).

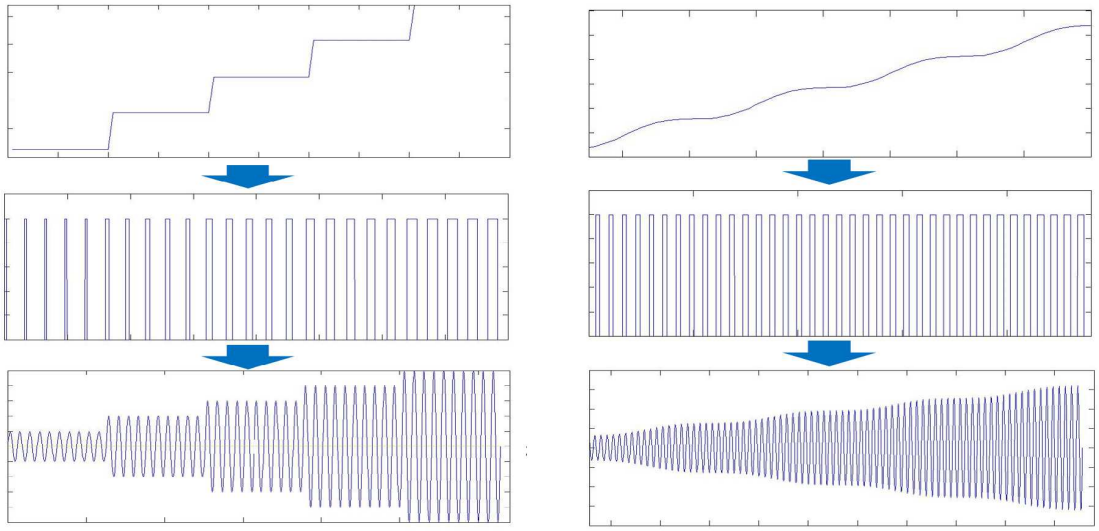

Fig. 5.7. Simulative output signals (Bottom) of the input driving voltage (Middle) determined from the sequence of duty cycles (Upper).

cover vibration up to 1 kHz, an FIR filter with low-pass characteristics blocking 1kHz or higher input components is what to design. We design a 20th order FIR filter

$$B(z) = \sum_{i=0}^{20} b_i z^{-i} \tag{5.6}$$

,whose coefficient is given by
$[b_0 \ \ldots \ b_{20}] = [0.0050 \quad 0.0070 \quad 0.0125 \quad 0.0215 \quad 0.0336 \quad 0.0478 \quad 0.0627 \quad 0.0768 \quad 0.0882 \quad 0.0957 \quad 0.0983 \quad 0.0957 \quad 0.0882 \quad 0.0768 \quad 0.0627 \quad 0.0478 \quad 0.0336 \quad 0.0215 \quad 0.0125 \quad 0.0070 \quad 0.0050]$.

Figure 5.8 shows the filtering result of $B(z)$ with a pulse train period of 20 samples. Output similar to sinusoidal signal is seen.

It should be noted that all of the filtering procedures are what are applied not directly

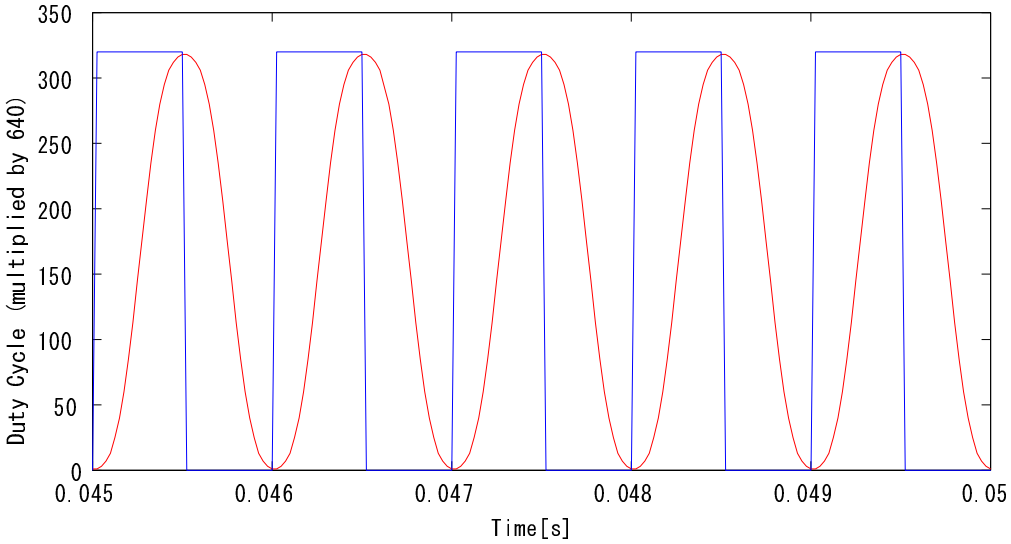

Fig. 5.8. Filtering the original sequence of duty cycle (blue) to smoother one (red).

to the output driving signals or ultrasound waveform, but to the value of discrete duty cycles of driving voltage pulses.

5.3 Experimental Verification

5.3.1 Numerical Simulation

We verified the validity of the proposed method in experiments. To begin with, a numerical evaluation of the filter $B(z)$ was operated. Figure 5.9 depicts the output waveforms for driving voltages of 40kHz whose duty cycle value takes 320 and 0 alternatively at 20 samples (resulting in 1 kHz switching of duty cycles).

The lower figure in 5.9 shows the output waveform with the duty cycle filtering. It is seen that the envelope in the lower figure is much closer to 1kHz sinusoidal wave compared to the upper figure indicating output without the duty cycle filtering. This difference can be seen much clearly in Figure 5.10, showing down-sampled waveforms of those in Figure 5.9. Those waveforms contain only audible components. As seen in the previous section, spiky noises are observed in the upper figure (non-filtered). Their power spectrum is shown in Figure 5.11. For the signal without filtering (blue line), periodic comb of 2kHz interval is seen. It is observed that it contains power in audible frequency ranges. On the other hand, this spiky comb is not found in the spectrum of the filtered output (red line), which is replaced with a sinc-function-like frequency structure. This sinc-function-like structure is what the author thinks indelible because it is derived from the intrinsic discontinuous property of the output envelope caused by pulse width modulations. The discontinuous envelope can be considered as a convolution of rectangle window and impulse series multiplied by desirable envelope of 1kHz bandwidth. Eventually it results in containing sinc-functional structure no matter how short the pulse period is. After all, the audible components are observed much reduced.

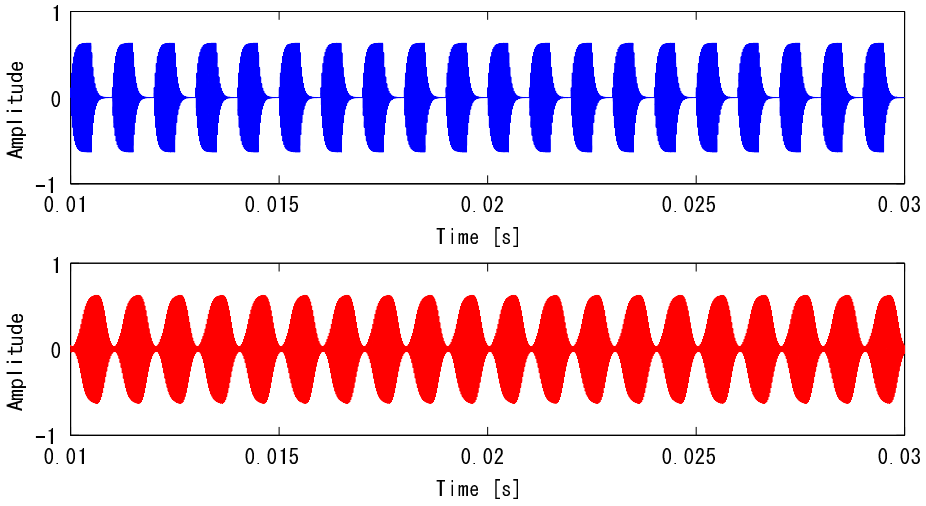

Fig. 5.9. Output waveform without (Upper, blue) and with (Lower, red) duty cycle filtering.

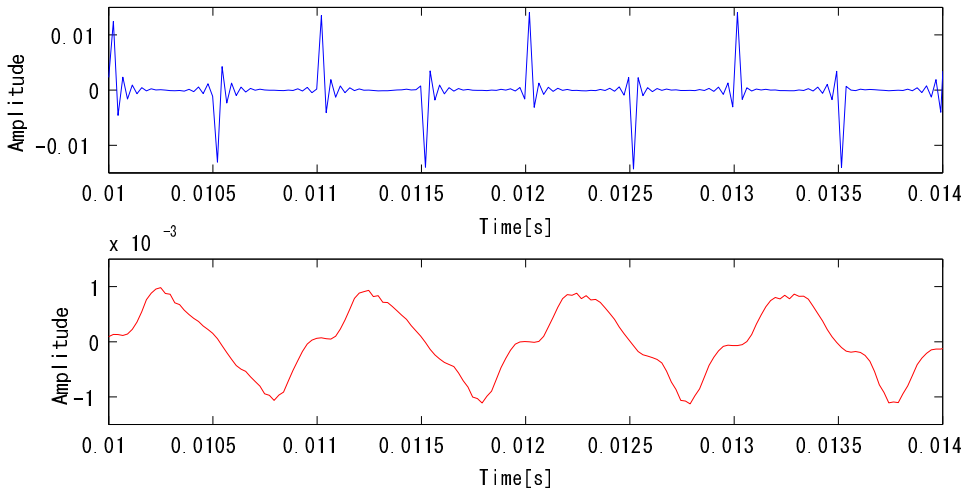

Fig. 5.10. Down-sampled waveforms of those in Figure 5.9. the line colors corresponds with Figure 5.9.

5.3.2 Evaluation using Real Recorded Signals

We operated another experiment using real recorded signals generated from the AUTD with the implementations described above. The incoming duty cycle signal from the MCU was set to 200 Hz sine wave. Sound amplitude was recorded at a focal point 10 cm apart from a single-unit AUTD using the same recording apparatus as in Chapter 4. The background noise was removed from recording signals. The recording was conducted for two implements: output signal with and without duty cycle filtering.

Figure 5.12 and 5.13 shows the spectrum of recorded signals. A periodic frequency structure due to the 200Hz component of input signal can be observed in both spectra. In comparison with the spectrum of the non-filtered signal, power of 5kHz or higher

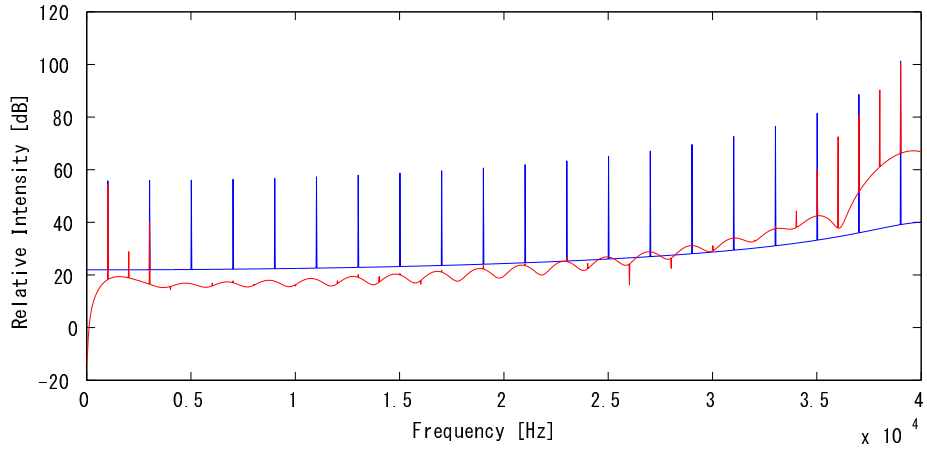

Fig. 5.11. Spectrum of waveforms in Figure 5.9. the line colors corresponds with Figure 5.9 and Figure 5.10.

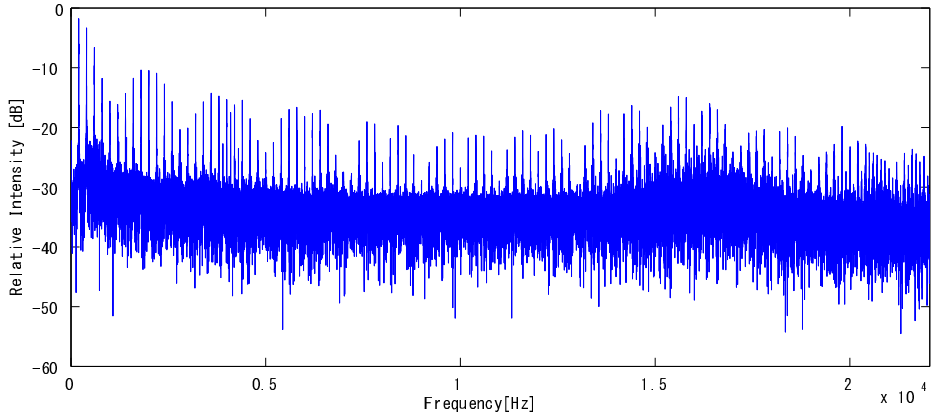

Fig. 5.12. Spectrum of recorded signal without duty cycle filtering.

frequency component drastically vanishes. As a result, the generated noises were heard being reduced.

5.4 Discussion

The experimental results have shown that the proposed duty cycle smoothing works to some degree. However, there are still many issues to be investigated further. The output power out of filtered driving voltages is attenuated as seen in temporal waveforms depicted in figures above. It is natural since the time-average of output ultrasound is reduced by filtering. On the other hand, human tactile modality has a low-pass characteristic at 1 kHz. By integrating these facts, one may think that it is possible that generated sensations to maintain its perceptive magnitude in spite of physical attenuation of ultrasound. However, the sensation conveyed by ultrasound focuses out of filtered voltages became tactually weaker. More quantitative investigation should be done for the clarification of tactile and auditory effects of the proposed filtering.

There is another important issue. The carrier ultrasound $p_c(t)$ is assumed to be a sinusoidal wave in this chapter. Howerever, nonlinear behavior of sound field such as described

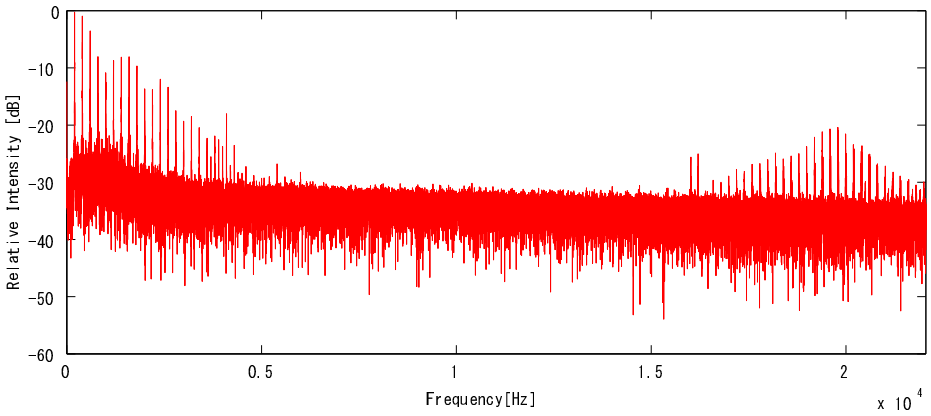

Fig. 5.13. Spectrum of recorded signal with duty cycle filtering.

above certainly has also much to do with the mechanism of noise generation. Since all of the noise generation mechanisms described above are only linear phenomena of sound, no nonlinear interactions of ultrasound are taken into account. For example, Because of high intensity of sound field, the wave distortion is caused as in Figure 3.1 in Chapter 3. The distorted waveforms contain complicated frequency structures which is much different from the original waveform of the transducers. There is a report that distorted sinusoidal output waves result in accompanying multiple harmonics, which causes audible noises by their interactive interferences. The more detailed modeling of the practical sound fields, the distortion of the output waveform should be taken into account. It is quite possible that further investigations considering nonlinear nature of intense sound field lead to more sophisticated solution to suppress audible noises to higher degree.

Chapter 6

Real World Systems

In this section, two real-world demonstration systems which make use of remote vibrotactile stimuli generated by AUTD are described. The second system is constructed in coorperation with Ishikawa Oku laboratory in the University of Tokyo. These two systems are for exhibitions in academic conferences and technological expositions. Many participants of those events have experienced both systems.

6.1 Tactile Projector

6.1.1 Overview

The author has proposed a passive ambient display system, which is specified in projecting visual images and ultrasound tactile vibration on the human body simultaneously. The aim of this system is to provide users with visual and touchable objects on their body surfaces.

Figure 6.1 is an image of the proposed system. The system consists of 4-unit AUTD and an video projector both mounted on the ceiling of the jig. Spatial calibration between projected images and ultrasound focus is done in advance. Several different virtual images with their individual appearance and corresponding vibrotactile sensation are projected in the same positions. The vibrations are extracted in the way described in Section 4.2. The projected objects are:

Fig. 6.1. Proposed 'Tactile Display.' unplugged. [63]

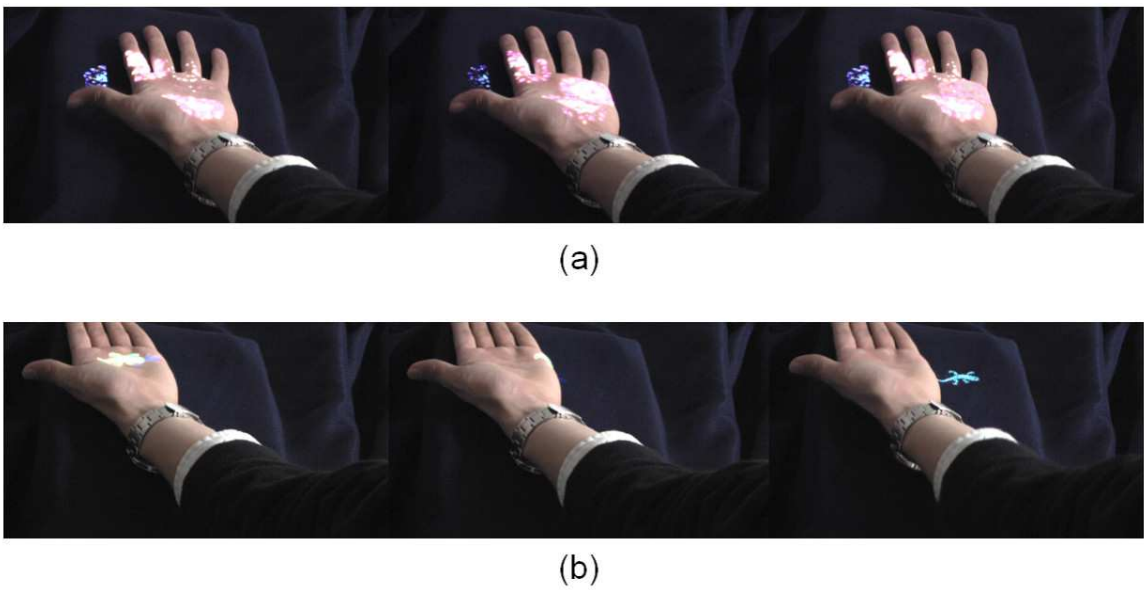

Fig. 6.2. Vibrotactile stimuli overlapped by projected images: (a) Fireworks, (b) a lizard creeping off the palm.

- Walking ladybug
- Crawling caterpillar
- Locomotive running on rails
- Rapturing fireworks
- Creeping lizard

As applications, amusement use such as movies which projects a part of the scene over our bodies like creeping worms is what the author reckons promising. We have also fabricated a larger system with feedbacks of user's body movement, which is discussed in the next section. This application is an example of taking advantages of the passive aspects of human tactile modalities. Merely projected images are not perceivable if we do not see them. By overlapping tactile stimuli, we can be aware of them tactually. In the system ultrasound tactile vibration is used since it does not violate the projected image visually. Thanks to the transparency of focused ultrasound, a natural visuo-tactile stimulation on ehe skin is successfully created.

6.1.2 Operation

Figure 6.2 shows the examples of projected objects. By adding vibrotactile sensations over projected images, virtual existence of the objects can be felt by users visually and tactually. Tactile stimuli on the objects are conveyed by a single ultrasound focal spot. Users are supposed to lay their hand with its palm upside on a stand placed under the device. The distance between each component to the user's palm is adjusted so that it fits the spatial calibration in advance. The system is a feed forward system, which projects images and tactile sensations whether users' hand is placed or not.

6.1.3 Demo at IEEE World Haptics Conference 2013

In IEEE WHC 2013, we performed a demo on the system 6.3. During the demo, projected objects were switched manually. Among of all five objects, the creeping lizard seemed to

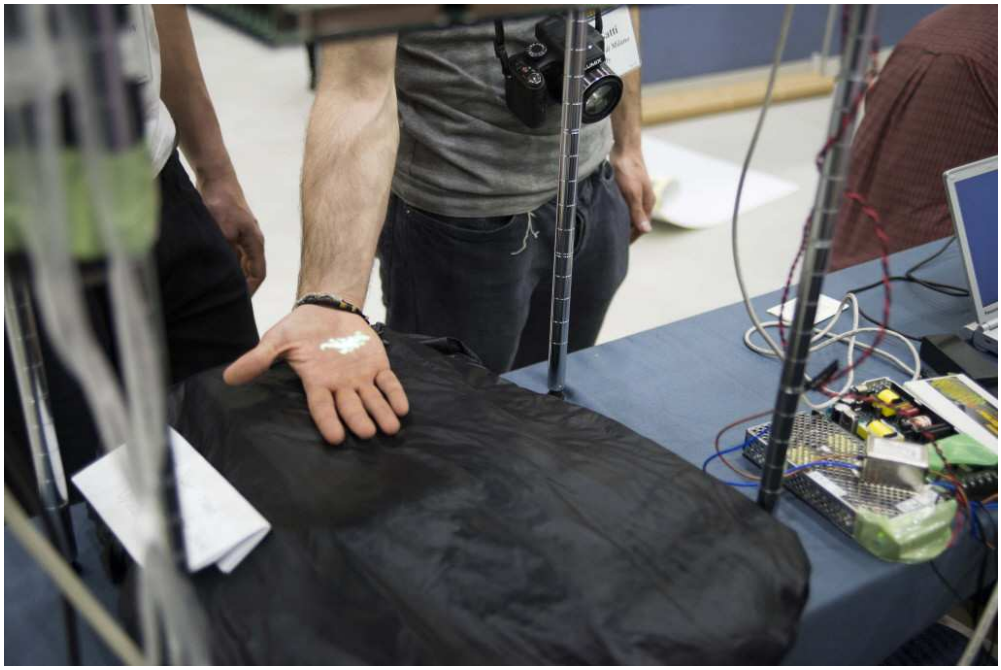

Fig. 6.3. Exhibition at the IEEE WHC 2013[63].

be the most popular and appealing to participants. Some female participants persistently refused to experience the system after she had been touched by the virtual lizard, which had made her surprised much. Its vibrotactile texture is created with a vibration with a soft impulse. Since this system physically pushes the object placed in the workspace, some participants found that a generic object such as a piece of paper is pressed by the radiation pressure and was surprised to see that. The physical 'weight' of the projected object is realized with this system, which might enhance the reality of the existence of them. The fireworks seemed to be the most easily perceptive since the primary frequency components in the vibrotactile stimuli was around 100∼200 Hz, where cutaneous sensation is perceived at its most sensitivity. For the crawling caterpillar, few participants said that it was realistic in touch sensation. The presented tactile sensation was 20Hz sinusoidal vibration, which gave us the feeling of being touched by something soft moving slowly. However it seemed that the tactile stimulus was not so intuitive that few users associated it to the projected image of caterpillar.

6.2 High-Speed Dynamic Information Environment

6.2.1 Overview

This is a corroborative work with Ishikawa-Oku laboratory in the University of Tokyo. A research group to which the author belongs has proposed the High-Speed Dynamic Information Environment. The concept of the research is to embed intellectual functions and informative display technology on the architecture or object existing in advance. This concept is in contrast to conventional personal computing systems or mobile gadgets with their entire functions implemented in themselves. The similar primitive conceptual portent of the research is partially seen in [45].

Nowadays, network communications have so widely been spread that it is possible for us to get connected to them by our mobile devices at a huge number of places. This

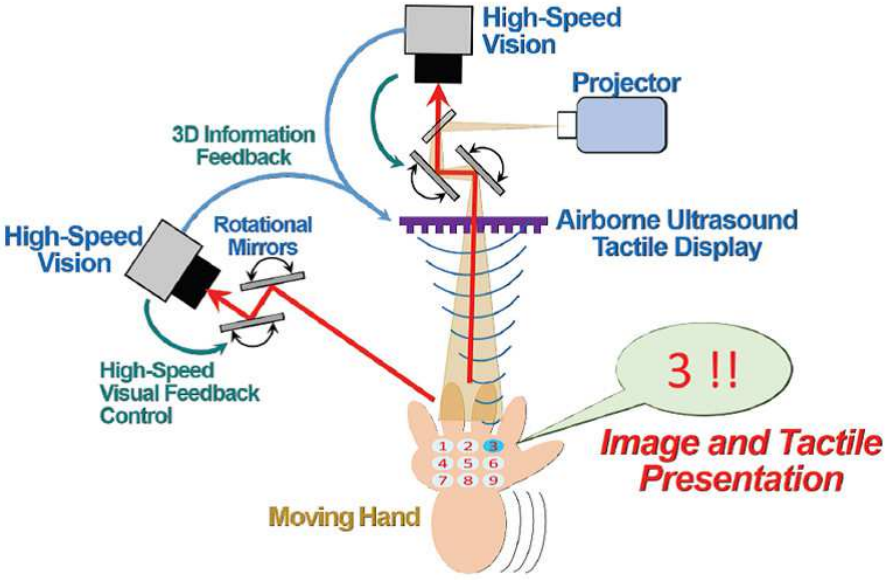

Fig. 6.4. Conceptual image of proposed system.

situation is completely different from that in 10 or 15 years ago when it was necessary to use a wired terminal or there were few places providing wireless networks. We can say that the network has been ubiquitous in this sense.

The author would like to emphasize that there is a great gap between current interfaces and networks in the way how they are operated. As mentioned above, interfaces are owned by users while the network has become ubiquitous and almost public. The term 'interface' literally stands for the margin of users and information provided by the network in this context. It converts the information into perceptive sensations such as vision or sound. Perhaps the premise that interface apparatus whose existence is recognizable are always necessary seems rational.

However, what if we can directly access to the information of networks with our sensory systems? What if the environment can provide us with the whole necessary information converted into perceptive sensations keeping us unconscious of any 'interfaces' around us? In the research we aim to realize such an interface which is in harmony with the way current network systems providing accesses, that is to say, a ubiquitous information display system connected with currently existing ubiquitous network systems.

There will be plenty of advantages over the current devices in the proposed framework. The following are examples. First, devices owned by individuals will be free from being under improper uses of malicious users. There will be no need for maintenance of personal devises. The communication compatibility between users is guaranteed since what each of individuals uses is embedded as the common system. In addition, users' movement is not constrained.

In the research, we have chosen visual and tactile modalities for information display. The reason is that tactile modalities can take attentions of users and make them ready to receive informations and a wide variety of informations is conveyed by visual modalities. Figure 6.4 is a conceptual image of the system. The system is designed so that they can project images and tactile stimuli on the human bodies tracking their movements with little latencies. With the system, human bodies and generic objects in the environments can be transformed into input devices with visual and tactile feedbacks. In other words,

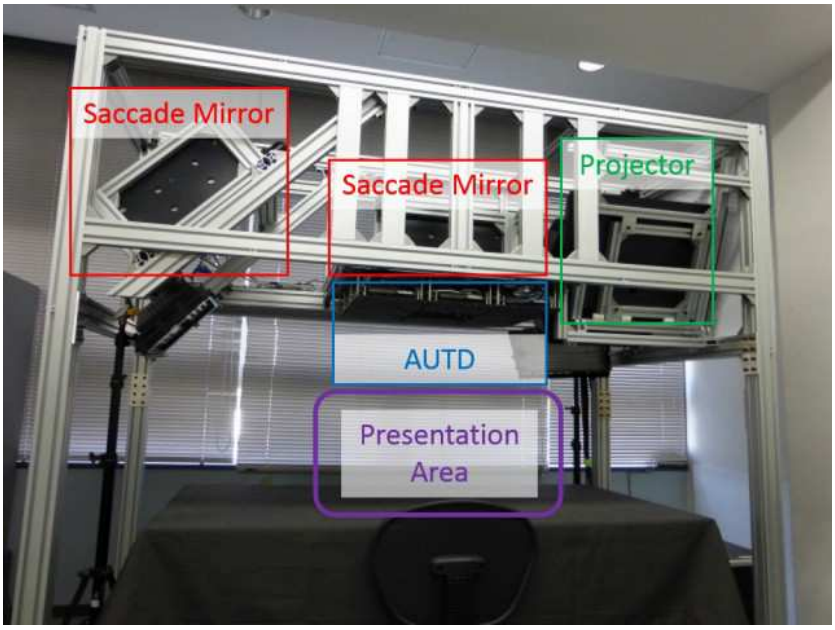

Fig. 6.5. fabricated system.

a vision and tactile projection mapping technology for moving objects at high speed has been developed in the research.

6.2.2 System Construction

The system is composed of AUTD and the high-speed dynamic projecting system[67] and a computer(Figure 6.5) [69][70]. Both components operate with little latencies due to their physical properties. Since AUTD is described in detail in the main article, there would be no need for further discussion. The image projecting system by Okumura et al.[67] uses multiple rotating mirror for controlling the angle of view and projection, which enables a high-speed tracking of a moving palm. High-speed vision sensors are employed to capture images necessary for mirror angle control in real time. Two object tracking units called Succade Mirror with high-speed image sensors of 512 pixels × 512 pixels resolution (FASTCAM MC2 by Photron) are operated for three-dimensional palm tracking in 500 fps. For the image projection a image projector with refresh rate of 120 Hz (SONY VPL-HW30ES) is used. The PC implemented in the system is Intel Xeon (CPU clock of 2.83GHz ,3GB Internal memory, 32bit Windows XP). The palm region detection is based on hue detection of input images.

6.2.3 Latency Test

Tactile latency of the system was tested for moving palms. The physical latency of the AUTD is about 4ms in current setup as demonstrated in the Section 4.1. The experiment was operated for evaluating the psycophysical effects of the latency when integrated into the proposed system.

Five subjects (all in twenties including one female subject) were asked to move their palm under the apparatus apart from the AUTD by 50 cm (Figure 6.7). During the experiment, an ultrasound spot of 30 Hz sinusoidal vibration is displayed on the center of the palm whose position is dynamically detected by the 500 Hz tracking system. There was

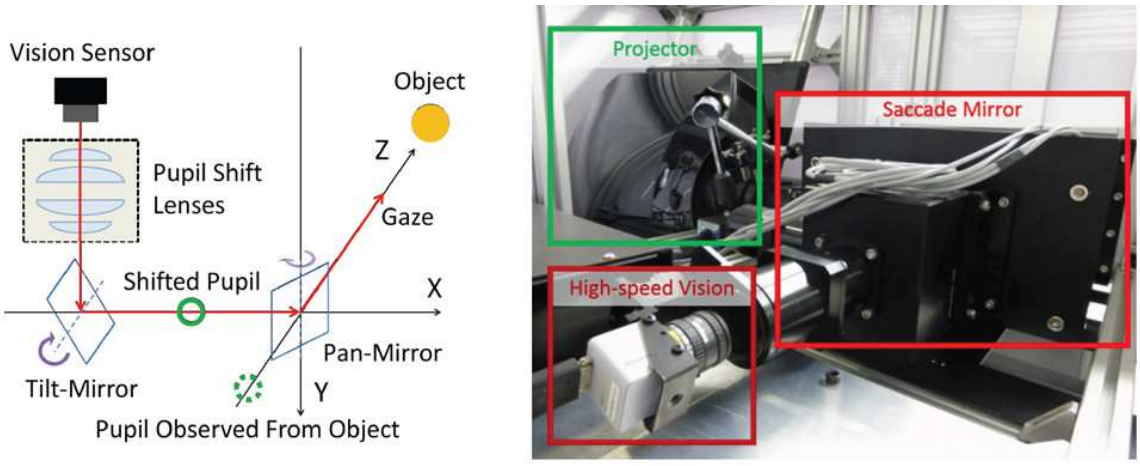

Fig. 6.6. high speed image projecting system with feedback embedded to the system[67].

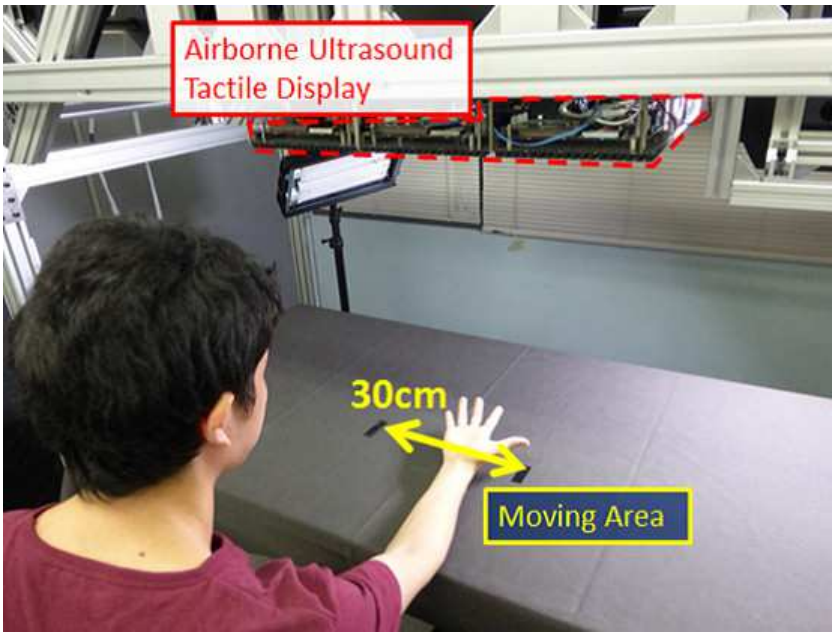

Fig. 6.7. Picture of latency test.

a direction on the movement of subjects' palm that it must move within the range of 30 cm laterally indicated by plastic tapes on the desk (this direction is defined as the x-axis in the experiment). The subjects were asked to accelerate movements of their palms, slowly in the beginning to their possible maximum velocities at the end of the experiment. When they felt that the displayed focal spot dropped off from their palms, they were supposed to inform it verbally. As a result, no one among the five subjects uttered that the focus dropped off. According to their movements recorded, the maximum period of them was about 0.2 s from end to end. The maximum velocity was abound 4 m/s. Figure 6.8 depicts the movement trajectory of a subject's palm in the x axis. It was shown that the 4 ms tactile delay in the system did not violate the palm tracking under the movement of this degree. However, the focal point did not stay at the same region during the experiment. The possible causes are errors of the palm detection algorithm and the actual tactile

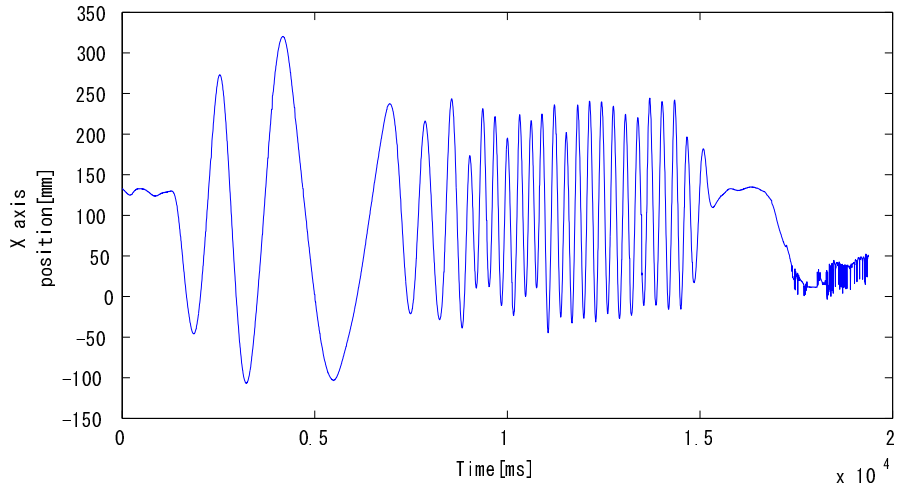

Fig. 6.8. Example of trajectory of palm movements.

delay caused by the AUTD. At the current experiment they are indistinguishable from each other. Spatial errors between images and ultrasound tactile stimuli projected on the palm by 1 cm have been shown tolerable[68].

Chapter 7

Conclusion

7.1 Summary of the Thesis

In the thesis, theoretical and experimental approaches for fully harnessing ultrasound as a medium for generating vibrotactile sensation are discussed.

In the beginning of the thesis, how a tactile display with acoustic radiation pressure contributes to more casual and versatile use of tactile applications as well as to tactile basic researches. The author demonstrated that there is a promising possibility for tactile modalities to be widely used in our daily lives in relation to its unique properties. Physical advantages of utilizing ultrasound as wave propagations are also presented with its high reproducibility, good spatial and temporal resolutions, low latencies and wide workspaces in accordance with past researches[11][4]. The recent tide in the human-interface researches are referred in the context of interfaces whose functions is expanded to or embedded in the surrounding environment. The author is convinced that this tide reflects the potential expectation for freeing users from being equipped with interface devices. The AUTD technique will be indispensable key for realizing that sort of future. At the end of Chapter 1, contributions of this thesis are clarified.

In Chapter 2, basics of human tactile perceptions investigated so far are briefly introduced. It is noted that the scope of the thesis is confined to cutaneous sensations, which is well understood by spatial and temporal behaviors of four mechanoreceptors. Their characteristics are described. In the middle of the chapter tactile display methods are categorized into three groups: mechanical-contact-based methods, electrical methods and remote stimulation methods. Each category of methods is described with several representative examples by stating what kind of tactile aspect is reconstructed by the method. Among mechanical methods, displays generating vibrotactile sensations are featured in its independent subsection. This is in intent to point up the fact that well-designed vibrations can generate a surprisingly realistic vibrotactile sensations on the skin. They even give us the feeling of inertial movement or weight variation inside the device. Moreover, a vibration at a single point inside a device can provide perception of spatial variation. This fact shows that by merely cutaneous vibrations a wide variety of phenomenon which physically do not exist can be percieved. In this sense the AUTD technique is absolutely promising since it can handle a wide variety of vibrations of a broad frequency band with spatial distributions. This chapter also clarifies the physical difference between ultrasound propagation and air flows. In the latter part of the chapter, examples of tactile applications currently in use and under investigation are presented. Their functions are classified into sensory substitutions, job trainings, amusement uses and tele-existance sensories. It should be noted that those examples are rather too professional or still for limited uses.

Chapter 3 begins with physical description of sound fields within the linear ranges. This is because in most of regions in space the linear description of sound field is applicable

in the situation of present study. The physical explanation of acoustic radiation pressure is given as a nonlinear behavior of sound with high intensity. Based on the physics above, presentation of methods to generate a localized ultrasound with designed amplitude distribution follows. The phased array technique is introduced as the first method for creating a single focal point of ultrasound which pushes the skin surfaces. This framework owes much to the researches by Iwamoto et al[11]. Estimation of resulting sound fields with a focal point is presented in accordance with the results in [4]. Numerically it is demonstrated that small array can not make a focus in far field without being blurred. A need for creating a large aperture array of transducer is implied with this result. The author proposed several methods which create three-dimensional distributions of acoustic radiation pressure by applying an inverse-problem framework to the problem. The most straight forward one of them is similar to [41], yet the author demonstrated it is applicable for three-dimensional distributions of transducers and presented radiation pressure. The goal of the proposed method is to generate a tactile object with spatially expanding volumes. It is expected that edges, faces or curvatures is reconstructed as perceivable tactile objects. The validity of this method is numerically verified through simulations. As a result, focusing performance in the depth direction by a planer array is shown to be much worse than that by arrays in which transducers are arranged in three-dimensional degree of spatial freedoms. Since it is also indicated that only a part of transducers actually contributes to pressure generation with this method, the author constructed another optimization method which intends to fully employ the power of all transducers and output pressure at regions where relatively intense output are required by tolerating errors in the other regions. As a result, for an appropriate set of tuning parameters, it is demonstrated more intense sound field can be generated by transducers with much more averaged output levels. The author also proposed a method which takes the resulting pressure directions into account for generating more realistic tactile object in the air. In the description of the method, the particle velocity distribution is what to be reconstructed by transducer arrays, which is supposed to be in relation to produced pressure directions. The procedure to create transducer output is derived and its performance is shown in the numerical experiments.

Chapter 4 gives detailed description of actually constructed vibrotactile display systems for generating a single vibrotactile spot in the air composed of newly fabricated multiple AUTD units. The whole signal transmission is presented along with the structures of signals containing information for controlling the position and amplitude of the focus. In the experiment, the system with 9 transducer units is verified to create a proper focusing comparative to the wavelength at the depth of 600mm away from the radiation plane, whose length is almost the same as the array size. The vibrotactile stimuli at the sampling rate of 2 kHz are generated by MCU based on the waveform stored in it and its internal clock. The amplitude control is done with the pulse width modulation of the driving voltage pulses imposed into transducers. A method to determine appropriate duty cycles of the pulses for the representation of desired vibrotactile sensations is derived. The chapter shows an example of recorded sound amplitudes and estimated radiation pressure waveform based on the pulse generating method above. Also, psycophysical experiments to estimate variation of the displayed vibrotactile stimuli were performed. As a consequence, it is confirmed that some of the displayed stimuli are easier to be distinguish then others. Temporal envelope of high carrier frequency seems more difficult to detect. The vibrotactile signals extracted from audio signals are also presented along with its subjective expression of tactile textures.

In chapter 5, a method for reducing the audible sound is proposed and validated along with revealing of what caused them. The mechanism of noise generation is explained by the spectrum deformation due to the amplitude modulation operated for vibrotactile

stimuli generation discussed in the Chapter 4. Based on this hypothesis the author has proposed to smooth the temporal change in a series of duty cycles of driving voltages on ultrasound transducers. It is numerically and experimentally validated that the proposed method works well in reducing audible sounds. Those results are followed by a discussion over its effects on produced tactile sensations and necessity of taking nonlinear acoustic behaviors into account.

In the chapter 6, actually constructed real world demonstration systems are described as practical examples to show the applicative possibility. They are 'Tactile Projector' and 'High-Speed Dynamic Information Environment' . The former provides users with a visual image and vibrotactile stimuli overlapping it on their palm with an image projector and AUTD. Users can visually and tactually feel the existence of the projected objects. The latter is constructed in cooperation with Ishikawa Oku Laboratory in the University of Tokyo, which is equipped with the high-speed tracking and projecting system developed in their laboratory. Owing to the tracking system, precise projection mapping on a human palm moving at the speed of our fastest waving is realized with no latency. The system can provide users with visible and touchable images projected on their palms. These two systems have been operated in exhibitions several times and many participants in academic conferences and technological expositions have experienced both systems. In the first chapter, the author mentioned that convenience of tactile displays in terms of having users get rid of wearing bulky devices is an intrinsic contribution for the prevalence of tactile applications. They are what the author thinks as possible examples of a brand new style of information display environments near future. These systems meet the needs above. Moreover, the High-Speed Dynamic Information Environment tolerates free and natural movements of users. This achievement can lead to a future where sports players receive tactile stimuli during the games and a lot of commuters are guided for the directions by remote tactile stimuli.

7.2 Future Perspective

The thesis presents some promising results in proposed and fabricated vibrotactile display system. In the following the author lists up possible future perspective.

7.2.1 Establishment of a Procedure for Producing High Fidelity Vibrotactile Sensation

One of the most intrinsic improvements in the present study could be the quantitative establishment of vibrotactile stimuli generation. In the thesis, how to create a desirable vibrotactile stimulus which reproduces realistic textures is not investigated in detail. The same problem remains in the spatial amplitude distribution generation methods above. In the thesis it is assumed that the ideal spatial and even temporal distribution of amplitude and particle velocity is given in advance. Nevertheless, since the relation from a physical phenomenon and its perception is the most intrinsic problem to be solved, the further research is supposed to handle this issue.

In the context of tactile reproduction system, no optimal sensor for the AUTD system has been proposed. Since the phenomenon utilized in the system is unique to it, there should be its inherent sensing framework capturing necessary and sufficient physical information for its reproduction via AUTDs. Researches on development of this AUTD-oriented sensor are another big issue for the study.

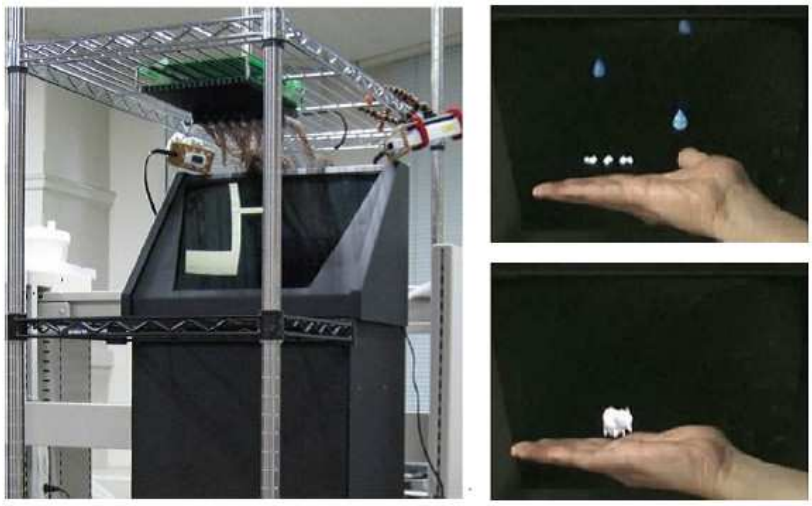

Fig. 7.1. Touchable Holography [61].

7.2.2 Vibrotactile Display with Time-Variant Stimuli based on Feedbacks of User Movements

The proposed system can be extended to the interactive system providing tactile feedback. In conference demonstration, Hoshi et al. performed the interactive system enabling users to touch floating images with tactile feedbacks [61] (Figure 7.1). It is true the system in provided with tactile feedback, it is merely repetitive vibrations displayed on the skin at the timing that the sensing system detects users' touching the images. To take the most advantages of the system, small physical latency, an interactive tactile model can be introduced for the sake of tactile quality enhancement. It can provide the feeling of touching invisible deforming object or objects with more complicated structure described with their internal states. A box with a rolling ball inside is an example of what this modification would enable to produce. An object of a large volume can also be reproduced by presenting three-dimensional tactile feedback at the point where the user skin is in touch with the object.

7.2.3 Real World Experiments

The further validation of methods described in the thesis is indispensable especially in real world operations. It is needless to say that they must be conducted by physical experiments, what is dealt with the thesis is human tactile sensation. Therefore physiological and psycophysical experiments with adequate setups are to be operated for the enforcement of the validity of the proposed system. Further implementation is required for realizing all the improvements above.

7.2.4 Use of other Nonlinear Nature of Sound Fields

What is described in the thesis among numerous nonlinear acoustic phenomena remains only a small part of them. There must be another nonlinear behaviors which can be harnessed in the proposed system. Further investigation and understanding over nonlinear acoustic fields will help applying new principles to aerial vibrotactile displays. For

example, reduction of the airflows with more shrewdly controlled sound field can help improving the fidelity of tactile stimuli since the air flow overlaps the vibrotactile focus with the current AUTD. As a different approach to this problem, Yoshino et al. has shown that a layered plate with slits blocks the airflow generated by the intense ultrasound spot while it lets propagating ultrasound pass through [71]. The air flow is not only something which is supposed to be terminated. By controlling the direction and velocity of it, tactile experience of the actual air flows with relatively small spatial distribution in regions apart from the AUTD, which is impossible for ordinaly air-jet systems to generate.

Bibliography

[1] C.E.Shannon, "A Mathematical Theory of Communication," Bell System Technical Journal, 27, pp. 379–423 & 623–656, 1948.
[2] M. Takahashi and H.Shinoda, "Large Aperture Airborne Ultrasound Tactile Display Using Distributed Array Units," Proc. of SICE Annual Conference, pp. 359–362, 2010.
[3] K. Hasegawa and H. Shinoda, "Aerial Display of Vibrotactile Sensation with High Spatial-Temporal Resolution using Large-Aperture Airborne Ultrasound Phased Array," Proc. of IEEE World Haptics Conference, 2013.
[4] T. Hoshi, et al., "Noncontact Tactile Display Based on Radiation Pressure of Airborne Ultrasound," IEEE Transactions on Haptics, vol. 3, no. 3, pp. 155–165, 2010.
[5] J. Awatani, "Studies on Acoustic Radiation Pressure. I(General Considerations)," Journal of the Acoustical Society of America, Vol. 27, pp. 278–281, 1955.
[6] P. J. J. Lamoreet et al., "Envelope Detection of Amplitude-Modulated High-Frequency Sinusoidal Signals by Skin Mechanoreceptors," Journal of the Acoustical Society of America, Vol. 79, No. 4, pp. 1082–1085, 1986.
[7] K. Hasegawa and H. Shinoda, "A Method for Distribution Control of Aerial Ultrasound Radiation Pressure for Remote Vibrotactile Display," Proc. of SICE Annual Conference, pp. 223–228, 2013.
[8] J. C. Gwilliam, A. Degirmenci, M. Bianchi and A. M. Okamura, "Design and Control of an Air-Jet Lump Display," IEEE Haptics Symposium , pp. 45-49, 2012.
[9] Y. Suzuki and M. Kobayashi. "Air Jet Driven Force Feedback in Virtual Reality," IEEE Conputer Graphics and Applications, Vol. 25, pp. 44–47, 2005.
[10] R. W. Lindeman, Y. Yanagida, H. Noma and K. Hosaka, "Wearable Vibrotactile Systems for Virtual Contact and Information Display," Virtual Reality, Vol.9, pp. 203-213, 2006.
[11] T. Iwamoto, M. Tatezono and H. Shinoda, "Non-contact Method for Producing Tactile Sensation using Airborne Ultrasound," Proc. of EuroHaptics, pp. 504–513, 2008.
[12] A. B. Vallbo and R. S. Johansson, "Properties of Cutaneous Mechanoreceptors in the Human Hand Related to Touch Sensation," Human Neurobiology 3, pp.3–14, 1984.
[13] N. Asamura, N. Yokoyama, H. Shinoda, "Selectively Stimulating Skin Receptors for Tactile Display," Computer Graphics and Applications, IEEE , vol.18, no.6, pp.32–37, 1998.
[14] T. Hirobe, S, Kuroki, K sato, T. Yoshida, K. Minamizawa, and S. Tachi, "Colorful touch palette," Proceedings of ACM SIGGRAPH 2010, Article No. 10, 2010.
[15] R. Sodhi, I. Poupyrev, M. Glisson, and A. Israr, "AIREAL: Interactive Tactile Experiences in Free Air," ACM Trans. Graph. 32, 4, Article 134, 2013.
[16] R. J. Mailloux, "Phase Array Antenna Handbook," New York: Artech House Inc., 1994.
[17] K. Hasegawa, N. Ono, S. Miyabe, and S. Sagayama,"Blind Estimation of Locations and Time Offsets for Distributed Recording Devices," Proc. of LVA/ICA, pp.57–64, 2010.
[18] K. Minamizawa, H. Kajimoto, N. Kawakami and S Tachi, "Wearable Haptic Display

to Present Gravity Sensation - Preliminary Observations and Device Design," Proc. of IEEE WorldHaptics, pp. 133-138, 2007.
[19] E. Hoggan, S. Brewster, and J. Johnston, "Investigating the Effectiveness of Tactile Feedback for Mobile Touchscreens," in Proc. of CHI, pp.1573–1582, 2008.
[20] S. J. Bolanowski , G. A. Gescheider , R. T. Verrillo and C. M. Checkosky "Four Channels Mediate the Mechanical Aspects of Touch," J. Acoust. Soc. Amer., vol. 84, pp.1680–1694, 1988.
[21] R. S. Johansson and A. B. Vallbo, "Tactile Sensory Coding in the Glabrous Skin of the Human Hand ," TINS, pp.27-32, 1983.
[22] C. R. Wagner, S. J. Lederman and R. D. Howe, "A tactile shape display using RC servomotors," Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2002. HAPTICS 2002. Proceedings. 10th Symposium on, pp.354–355, 2002.
[23] G. Moy, C. Wagner, R. S. Fearing, "A compliant tactile display for teletaction," Proceedings. of IEEE ICRA., pp.3409–3415 vol.4, 2000.
[24] T. Nara, M. Takasaki, T. Maeda, T. Higuchi, S. Ando, and S. Tachi, "Surface acoustic wave tactile display," Computer Graphics and Applications, IEEE , vol.21, no.6, pp.56–63, 2001.
[25] K. Minamizawa, S. Fukamachi, H. Kajimoto, N. Kawakami, S. Tachi, "Wearable Haptic Display to present Virtual Mass Sensation," ACM SIGGRAPH, Sketches, 2007.
[26] Q. Wang and V. Hayward, "Compact, Portable, Modular, High-performance, Distributed Tactile Transducer Device Based on Lateral Skin Deformation," Proc. 14th Symposium on Haptic Interfaces For Virtual Environment And Teleoperator Systems IEEE VR, pp. 67–72, 2006.
[27] G. L. Aiello, "Multidimensional electrocutaneous stimulation," IEEE Trans. Rehabil. Eng., Vol. 6, No. 1, pp. 95–101, 1998.
[28] R. M. Strong and D. Troxel, "An Electrotactile Display," Man-Machine Systems, IEEE Transactions on , vol.11, no.1, pp.72–79, 1970.
[29] H. Kajimoto, N. Kawakami and S. Tachi, "Electro-Tactile Display with Tactile Primary Color Approach," Int. Conf. on Intelligent Robots and Systems(IROS), 2004. .
[30] H. Kajimoto, "Electro-tactile Display with Real-time Impedance Feedback," Proc. of EuroHaptics, pp. 184–188, 2010.
[31] H. Y. Yao and V. Hayward, "An Experiment on Length Perception with a Virtual Rolling Stone," in Proc. EuroHaptics Int. Conf., pp. 275–278, 2006.
[32] K. Minamizawa, Y. Kakehi, M. Nakatani, S. Mihara and S. Tachi, "TECHTILE Toolkit," ACM SIGGRAPH, Emerging Technologies, 2012.
[33] H. Culbertson, J. Unwin, B. E. Goodman and K. J. Kuchenbecker, "Generating haptic texture models from unconstrained tool-surface interactions," IEEE World Haptics Conference, pp.295–300, 2013.
[34] J. G. Linvill and J. C. Bliss, "A Direct Translation Reading Aid for the Blind," Proceedings of the IEEE, Vol. 54, No. 1, pp. 40–51, 1966.
[35] L. H. Goldish and H. E. Taylor, "The Optacon: a valuable device for blind persons," New Outlook for the Blind, 68, pp. 49–56, 1974.
[36] Laerdal Medical. http://www.laerdal.com/
[37] MAKO Surgical Corp. http://www.makosurgical.com/
[38] C. L. Fernando, M. Furukawa, T. Kurogi, S. Kamuro, K. Sato, K. Minamizawa and S. Tachi, "Design of TELESAR V for Transferring Bodily Consciousness in Telexistence," Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2012), pp.5112–5118, 2012.
[39] T. Sakurai, H. Shinoda and M. Konyo, "Sharp Tactile Sensation using Superposition

of Vibrotactile Stimuli in Different Phases," Proc. IEEE World Haptics Conference 2013, Poster, pp.235–240, 2013.
[40] T. Sakurai, M. Konyo and S. Tadokoro, "Presenting Sharp Surface Shapes Using Overlapped Vibrotactile Stimuli," Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3300 - 3307, 2012.
[41] L. R. Gavrilov, "The possibility of generating focal regions of complex configuration in application to the problems of stimulation of human receptor structures by focused ultrasound," Acoust. Phys. 54, 2, pp. 269–278, 2008.
[42] Y. Ueda, Y. Mizusina, K. Minamizawa and S. Tachi, "Polymer-clay haptic display to present spacial vibro-tactile sensation," Proc. IEEE World Haptics Conference, Demo, 2013.
[43] The 4-D Experience in Vancouver Aquarium. http://www.vanaqua.org/experience/shows/4-d-theatre-shows
[44] C. Harrison, H. Benko and A. D. Wilson, "OmniTouch: Wearable Multitouch Interaction Everywhere," In Proceedings of the 24th Annual ACM Symposium on User interface Software and Technology (UIST). ACM, NY. 441-450, 2011.
[45] A. Zerroug, A. Cassinelli, and M. Ishikawa, "Invoked computing: Spatial audio and video AR invoked through miming," Proceedings of Virtual Reality International Conference, LAVAL VIRTUAL, pp.31–32, 2011.
[46] Y. Makino and H. Shinoda, "A whole palm tactile display using suction pressure," In IEEE Int. Conf. on Robotics and Automation, pp. 1524–1529, 2004.
[47] Y. Makino and H. Shinoda, "Selective Stimulation to Superficial Mechanoreceptors by Temporal Control of Suction Pressure," In World Haptics Conference, pp. 229–234, 2005.
[48] L. B. Porquis, M. Konyo, N. Nagaya and S. Tadokoro, "Multi-contact Vacuum-Driven Tactile Display for Representing Force Vectors Applied on Grasped Objects," Haptics:Perception,Devices,Mobility, and Communication International Conference, EuroHaptics, pp. 218-221, Volume II, 2012
[49] E. T. Enikov and K. V. Lazarov, "Micro-mechanical switch array for meso-scale actuation," Sensors and Actuators A, Vol. 121, No. 1, pp. 282–293, 2005.
[50] H. Shinoda, H. Chigusa, and Y. Makino, "Flexible Tactile Sensor Skin Using Wireless Sensor Elements Coupled with 2D Microwaves," Journal of Robotics and Mechatronics, Vol.22 No.6, pp.784–789, 2010.
[51] H. Shinoda, N. Asamura, M. Hakozaki and X. Wang, "Two-Dimensional Signal Transmission Technology for Robotics," Proc. of IEEE Int. Conf. on Robotics and Automation, pp.3207–3212, 2003.
[52] H. Shinoda, N. Asamura, T. Yuasa, M. Hakozaki, X. Wang, H. Itai, Y. Makino, and A. Okada, " Two-Dimensional Communication Technology Inspired by Robot Skin," Proc. IEEE TExCRA (Technical Exhibition Based Conf. on Robotics and Automation), pp.99–100, 2004.
[53] H. Kajimoto, Y. Kanno and S. Tachi, "Forehead Electro-tactile Display for Vision Substitution," Proc. EuroHaptics, pp. 75–79, 2006.
[54] J. C. Craig, "vibrotactile pattern perception: Extraordinary observers," Science. 22. No. 196(4288), pp.450–452, 1977.
[55] C. Y. Nolan and C. J. Kederis, "Perceptual factors in Braille word recognition," American Foultdation for the Blind Research Series, No.20, 1969.
[56] A. P. Grunwald, "A Braille-Reading Machine," Science. 7. No.154(3745), pp.144–146, 1966.
[57] J. M. Pickett and P. H. Pickett, "Communication of speech sounds by a tactual vocoder," Journal of Speech and Hearing Research , vol. 6, pp. 207–222, 1963.
[58] Nintendo Wii. http://www.nintendo.co.jp/wii/

[59] S. Tachi, K. Tanie, K. Komoriya and M. Kaneko, "Tele-existence (I): Design and Evaluation of a Visual Display with Sensation of Presence," RoManSy 84 The Fifth CISM-IFToMM Symposium, pp. 206–215, 1984.
[60] T. Kamakura, Fundamentals of Nonlinear Acoustics, Aichi Shuppan, Tokyo, 1996 (in Japanese).
[61] T. Hoshi, M. Takahashi, K. Nakatsuma, and H. Shinoda,"Touchable Holography," ACM SIGGRAPH, Emarging Technologies, 2009.
[62] T. Hasegawa, T. Kido, T. Iizuka, and C. Matsuoka, "A general theory of Rayleigh and Langevin radiation pressures," Acoustical Science and Technology, Vol. 21, No. 3, pp. 145–152, 2000.
[63] K. Hasegawa, Y. Monnai, M. Fujiwara, K. Yoshino and H. Shinoda, "An Aerial Vibrotactile Display with Floating Visual Images," IEEE World Haptics Conference, Demo, 2013.
[64] T. Iwamoto and H. Shinoda, "Two-dimensional Scanning Tactile Display Using Ultrasound Radiation Pressure," Proc. Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems (IEEE Haptics Symp., pp. 57–61, 2006.
[65] H. E. Bass, L. C. Sutherland, A. J. Zuckerwar, D. T. Blackstock, and D. M. Hester, "Atmospheric Absorption of Sound: Further Developments," J. Acoustical Soc. Am., vol. 97, pp. 680–683, 1995.
[66] D. A. Webster and D. T. Blackstock, "Finite-amplitude Saturation of Plane Sound Waves in Air," J. Acoust. Soc. Am. 62, pp. 518–523, 1977.
[67] K. Okumura, H. Oku and M. Ishikawa, "Lumipen: Projection-based Mixed Reality for Dynamic Objects," IEEE International Conference on Multimedia and Expo (ICME), pp.699–704, 2012.
[68] K. Yoshino, K. Hasegawa and H. Shinoda, "Measuring Visio-Tactile threshold for Visio-Tactile Projector," Proc. SICE Annual Conference, pp.1996–2000, 2012.
[69] K. Hasegawa, T. Sueishi, K. Okumura, M. Oku, M. Ishikawa and H. Shinoda, "Superposition of the Airborne Ultrasound Vibrotactile Stimuli on Visual Image Projected on Human Bodies Moving at High Speed," the 18th Conference of VRSJ, pp. 307–310, 2013 (in Japanese).
[70] T. Sueishi, K. Hasegawa, K. Okumura, M. Oku, M. Ishikawa and H. Shinoda, "Image and Tactile Presentation System for Moving Objects by Stereo Tracking Using High-speed Gaze Controllers by Rotational Mirrors," the 18th Conference of VRSJ, pp. 594–597, 2013 (in Japanese).
[71] K. Yoshino, H. Shinoda, "Visio-Acoustic Screen for Contactless Touch Interface with Tactile Sensation," Proc. IEEE World Haptics Conference 2013, pp.419-423, 2013.

Appendix A

List of Publications

Journal Articles

1. 長谷川圭介, 篠田裕之, "Tactile Projector:振幅変調空中超音波照射による多彩な触感を伴った映像の投影装置," 日本バーチャルリアリティ学会論文誌 (submitted, in Japanese)
2. 末石智大, 長谷川圭介, 奥村光平, 奥寛雅, 篠田裕之, 石川正俊, "空中超音波触覚ディスプレイ・カメラ系による高速ダイナミック情報環境とその校正手法," 日本バーチャルリアリティ学会論文誌 (submitted, in Japanese)

Referred Conference Papers

1. K. Hasegawa and H. Shinoda, "A Method for Distribution Control of Aerial Ultrasound Radiation Pressure for Remote Vibrotactile Display," Proc. of SICE Annual Conference, pp. 223–228, 2013.
2. K. Hasegawa and H. Shinoda, "Aerial Display of Vibrotactile Sensation with High Spatial-Temporal Resolution using Large-Aperture Airborne Ultrasound Phased Array," Proc. of IEEE World Haptics Conference, 2013.
3. K. Hasegawa and H. Shinoda: Dynamic Range Enhancement of Airborne Ultrasound Tactile Display, Proc. SICE Annual Conference 2012, pp.280-283, August, 2012.
4. K. Yoshino, K. Hasegawa and H. Shinoda, "Measuring Visio-Tactile threshold for Visio-Tactile Projector," Proc. SICE Annual Conference, pp.1996–2000, 2012.
5. K. Hasegawa, N. Ono, S. Miyabe, and S. Sagayama, "Blind Estimation of Locations and Time Offsets for Distributed Recording Devices," Proc. of LVA/ICA, pp.57–64, 2010.
6. K. Hasegawa, M. Kanno and S. Hara, "Algebraic Computation of the Stability Radius of a Multilinear Polynomial," Proc, ICCAS-SICE, pp.3531–3536, 2009.

Exhibitions

1. K. Hasegawa, Y. Monnai, M. Fujiwara, K. Yoshino and H. Shinoda, "An Aerial Vibrotactile Display with Floating Visual Images," IEEE World Haptics Conference, Demo, 2013.

Domestic Presentations

1. 長谷川 圭介，篠田 裕之，“空中超音波焦点形成による非接触触覚提示における反射壁を用いた気流の抑制,” 第 14 回 SICE SI 部門講演会論文集, pp. 2432-2434, 2013.
2. 吉岡 基，長谷川 圭介，篠田 裕之、“空中超音波触覚提示における駆動電圧パルス幅制御による可聴音の抑制,” 第 14 回 SICE SI 部門講演会論文集, pp. 2427-2431, 2013.
3. 末石智大，長谷川圭介，奥村光平，奥寛雅，篠田裕之，石川正俊, “駆動鏡面式高速視線制御を用いたステレオトラッキングによる動的対象への視触覚提示システム,” 第 18 回日本バーチャルリアリティ学会大会，pp. 594–597, 2013.
4. 長谷川 圭介，末石 智大，奥村 光平，奥 寛雅，石川 正俊，篠田 裕之, “運動する人体上へ高速追従する映像投影への空中超音波触覚の重畳,” 第 18 回日本バーチャルリアリティ学会大会，pp. 307–310, 2013.
5. 長谷川圭介, 篠田裕之, “超音波の三次元音圧分布制御による非接触触覚提示,” ロボティクス・メカトロニクス講演会 (ROBOMEC2013), 2013.
6. 長谷川圭介, 篠田裕之, “近接配置した複数アレイの協調動作による大開口空中超音波触覚ディスプレイの性能評価,” 第 13 回 SICE SI 部門講演会論文集, pp. 2244–2247, 2012.
7. 長谷川圭介，吉野数馬, 藤原正浩, 中妻啓, 篠田裕之, “Visuo-Tactile Projector:映像と触覚の同時投影装置の提案,” 第 17 回日本バーチャルリアリティ学会大会, pp.359–362, 神奈川, 2012.
8. 長谷川圭介, 篠田裕之, “空中超音波による把持物体を介した位置情報提示,” 第 12 回 SICE SI 部門講演会論文集, pp. 2244–2247, 2011.
9. 長谷川圭介, 篠田裕之 “把持物体への空中超音波照射による触覚提示,” 第 16 回日本バーチャルリアリティ学会大会, pp. 730–731, 2011.
10. 長谷川圭介, 小野 順貴, 宮部滋樹, 嵯峨山 茂樹, “信頼度重みづけを考慮した非同期録音のブラインドアラインメント,” 日本音響学会秋季研究発表会講演集, pp.591–592, 2010.
11. 長谷川 圭介, 宮部 滋樹, 小野 順貴, 嵯峨山 茂樹, “分散型マイクロホンアレーによる非同期録音信号のブラインドアラインメントの実環境評価,” 電子情報通信学会総合大会講演論文集, p.150, 2010.

Honors, Grants and Awards

1. Keisuke Hasegawa, Young Author’s Award SICE Annual Conference 2013, 2013.
2. Keisuke Hasegawa and Hiroyuki Shinoda, Finalist, WHC 2013 Best Paper Award, IEEE World Haptics Conference 2013, 2013.
3. Keisuke Hasegawa, WHC 2013 Best Student Paper Award, IEEE World Haptics Conference 2013, 2013.
4. 長谷川圭介, 原総合知的通信システム基金渡航助成, Proc. LVA/ICA, 2010.