

atMOS: Self Expression Movie Generating System for 3G Mobile Communication

Satoru Tokuhsa *, Taku Kotabe *, Masa Inakage**

Keio University *Graduate School of Media and Governance

**Faculty of Environmental Information

5322 , Endo , Fujisawa , Kanagawa , 252-8520 , Japan

+81 466 49 3545

{dk , taku , inakage}@imgl.sfc.keio.ac.jp

Abstract

This research focuses on movies as a new communication tool, and proposes a self expression movie generating system “atMOS”. The purpose of this research is to encourage mobile movie communication. Therefore, this system adopts the concept of exchangeability, expressivity and reproducibility for mobile movie contents.

In “atMOS”, based on the sub-concept of “self packaging movie”, users can make an original promotion movie of themselves which can be viewed on 3G cellular phones. Users can enjoy the original movie not only by themselves but also with other people by exchanging these movies as a communication tool, using the cellular phone.

Keywords

Media Design, Mobile, Movie, Cellular Phone, 3GPP
Media Communication, Performance, Sensing,

1 Introduction

Presently, the development of specs, services and contents for cellular phones are amazing.

As for the specs, sophisticated features are inevitable, and camera functions have become the standard for cellular phones.[1][2][3][4] Services are also enriched at present. For example, a debit service using digital money through the cellular phones [5], real-time route searching service using GPS and an acceleration sensor [6] is already implemented. As for the contents, automatic distribution service of movies is

brought in.[7] Some of these services can be used not only in Japan, but also in other countries. It is not only Japanese cellular phone manufacturers that make these advanced cellular phones. For example, NOKIA, Finland’s cellular phone manufacturer, also makes cellular phones with advanced design and function.

As stated above, technical development and development in services and contents for cellular phones are high-flying. From a research point of view, many researches have been presented in academic conferences like CHI, Mobile HCI, Ubicomp and Pervasive Computing, on cellular phones as a mobile information terminal. These researches can be classified broadly into 3 categories. First is the research of the interface of cellular phones. Research comparing and reviewing the menu designs [8], and research of command input method of text by 12 buttons aimed at beginners [9] can be stated as an example of this category. Second is the research of cellular phones as a information input device for a certain content service. MIVE [10] which is a system that a user uses a cellular phone to express their own feelings and reviews of a movie on a large screen while they are watching it is one example. A video mixing system [11] which can be operated by multiple users through cellular phones is another example of this category. Third is the Context-aware content service which uses the positional information and the network. For example, Noriyuki Ueda’s GIS with cellular phone +WebGis [12] create a virtual city by collaborating the GPS positional information and images taken by cellular phones. The Re: living Map [13] is another content which uses the GPS and camera function of cellular phones, but enables to re-experience and share a place where the user has walked through.

On the other hand, there are also many researches done from the sociological point of view. "Perpetual Contact" [14] has verified theoretically and practically the comprehensive survey and social interaction of cellular phones in countries all over the world, for example USA, France, Holland, Finland and Italy, based on the reports and researches of each country's researchers.

In terms of cellular phones as a communication media, the direct communication method using the present cellular phones can be classified into 4 ways, voice sound, text, still images and motion images. Therefore, direct communication by voice, communication by mail, communication by sending still images that has been taken and processed by users or communication by sending motion pictures taken by the users can be raised as these methods. Real-time motion image communication by cellular phones with picture phone functions can be categorized as a motion image communication.

Before stating the characteristics of mobile movies as a communication media, the main proposal of this research, the characteristics of image communication must be mentioned. Print Club [15] is a representative of analog image communication. Print Club is a system that creates an original photo sticker, combining the frames, stamps and scribbles by the user. The main purpose of this sticker is to express oneself, and also to exchange with other people. This stands as a media communication. An enjoyment of expressing oneself as intended can be attained and the sticker has a factor that will remind the situation just as photo picture does.

This analog image communication Print Club, which holds expressivity, exchangeability and reproducibility, was extended by the appearance of cellular phones with camera functions. A service "@Sha-mail" [16] that started in 2001 by J-Phone (currently Vodafone) enabled people not only to collect images they took with their camera functions on the cellular phones but also to send these images by the mail functions of the cellular phones. The convenience of being able to send and receive images, which was taken by the users needs, attached to the mail, made the exchangeability of Print Club to an intense expansion. Functions for picture frames and paints have appeared in some cellular phones, and expressivity

is advancing. NOKIA has put Medallion [17], a device which the user can wear one's favorite image, on the market. This can receive infrared data which enables users to wear others images, and presents another aspect to image communication.

Therefore, the basis of image communication in cellular phones lies on the concept of expressivity, exchangeability and reproducibility. Likewise, contents with expressivity, exchangeability and reproducibility are adequate as a communication media for movie communication in 3G cellular phones. At the same time, an expansion of the attributes associated with contents itself must be done as the media expands from still images to motion images. As it stands now, mobile movie communication is still at an early stage of development, The reason for the lack of positive contents is considered to be from the lack of these aspects. Consequently, we developed a self expression movie generating system which promotes mobile movie communication based on the structured concept model of expressivity, exchangeability and reproducibility.

2 Concept

The concept of the self expression movie generating system, which encourages mobile movie communication, can be summarized in the following 3 points:

- exchangeability: contents premised to be exchanged.
- expressivity: contents full of expressiveness
- reproducibility: contents that revive the users' feeling

For the exchangeability, in order to exchange, there is a prerequisite that other people wants the object. This prerequisite is affected by the rarity and originality. Limited versions of trading cards and figures have a rarity value. By giving rarity and originality in the contents, the exchange of the object can be encouraged.

For the expressivity, the expressiveness of movies is important. Making use of the new attribution centered on time, and to be more expressive than that of a still image is inevitable.

For the reproducibility, in this term, it can be defined as the reproduction of the users' feelings and senses by using this

content. The users' feelings and senses can be defined as the reminding of the intention and the memory which is the reproduction of body sensation, a new attribution in addition to the attribution held by the image itself. In order to reproduce the user's senses, the movement of the user must be strictly sensed (systematic reproduction of body sensation), directly connected to the expression and must be an enjoyable expression, just as the user intended (reproduction of intension). Based on this process, the system's aim is to reproduce not only the users' memory but also the body sensation attained by this content when the movie is seen.

Naturally, these three concepts refer to each other. In order to meet the rarity and originality of exchangeability, the reproducibility and expressivity is highly concerned. The cycle can be described that by adopting the reproducibility and the expressivity in the system, the exchangeability of the contents rises.

The self expression movie generating system, which adopts these three concepts to encourage mobile movie communication, is the proposal of this research "atMOS". "atMOS" aims to encourage the communication through a cycle of creation, acquisition and exchange, as seen in Figure 1. The main uncton of "atMOS" as a movie generating system is represented in Creation SelfExpression in Figure 1.

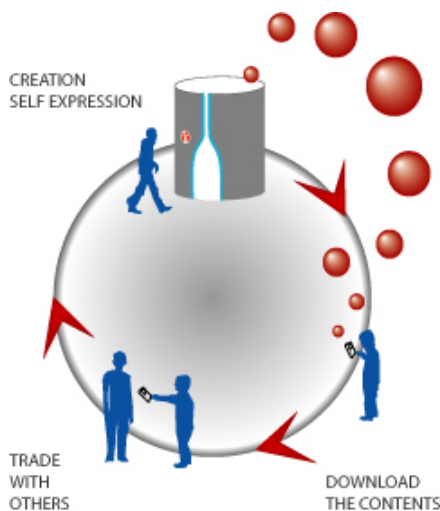


Figure 1 : atMOS Cycle

3 atMOS

"atMOS" (Figure 2) is a system that creates a promotion movie of the user, able to be viewed on 3G cellular phones, by simple actions and is based on a sub-concept of "self packaging movie".

This chapter describes in detail the system of atMOS from the expressivity, reproducibility and exchangeability point of view, together with the sub-concept "self packaging movie".



Figure 2 : atMOS

3.1 Self Packaging Movie

atMOS uses a sensing system to sense the motive actions of the user. Sound and image processing is done based on the digitalized data of the user's motive action. To process images, the image of the user, taken by the embedded digital video camera inside the chassis, is used as the main material. Users can download the result movie into their cellular phones. By saving the movie in their cellular phones, users can enjoy it not only by themselves but also by exchanging them with their friends. In terms of making sound and image processing based on the movement of oneself, and making oneself a material for a movie, it is packaging oneself as a content. Therefore, it is a "self packaging movie".

3.2 System Design

The system flow of atMOS is shown in Figure 3. It can be broadly categorized into 4 sections, as below.

- (a) Performance Section
- (b) Real-Time Effect Section
- (c) Rendering Section
- (d) Communication Section

Sensing of the users' performance is done in (a), and the data is digitalized and sent to (b). (b) allocates the data to the image and sound effects, and sends it to the monitor, speakers and (c) after processing. (c) gathers the movie data and encodes it to a format that can be viewed on cellular phones. (d) uploads the encoded movie to the server and sends to the user a mail showing the URL to download this movie.

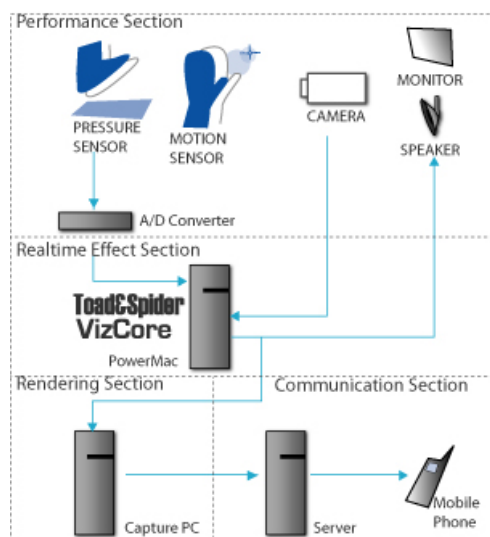


Figure 3 : atMOS System Flow

In the next place is how the three concepts, exchangeability, expressivity and reproducibility are implemented in this system of atMOS, made by this flow.

As for the development, Max/MSP and Jitter by Cycling(74) is used. The specs for the processing computers are PowerMac G4 (CPU Dual1.25GHz) when exhibiting for the first time at SIGGRAPH2003 [18], and PowerMacG5(CPU Dual2.0GHz) when exhibiting for the second time at Japan Media Art Festival [19]. These computers were used for (a) and (b), while ThinkPad31 was used for the capturing and rendering, (c) and (d).

3.2.1 Performance Section

The performance section senses the users' performance and makes the users' action into data. Here, the objective was to realize a high reproducibility of the users' motive action. This was done by sampling the movement of the user in a condition close to free hand, in order to reduce the factor that obstructs the movement.

atMOS uses a digital video camera and LED pointers to sense the body movements of the user. The movie, taken by the digital camera inside the chassis, is processed into a reversal image and screened on the monitor outside the chassis. (Figure 4)

The result of the image processing appears real-time on the monitor. Users can see which effect is linked to which movement, so they are able to play while directly checking the movie. This enhances the reproducibility of intention.

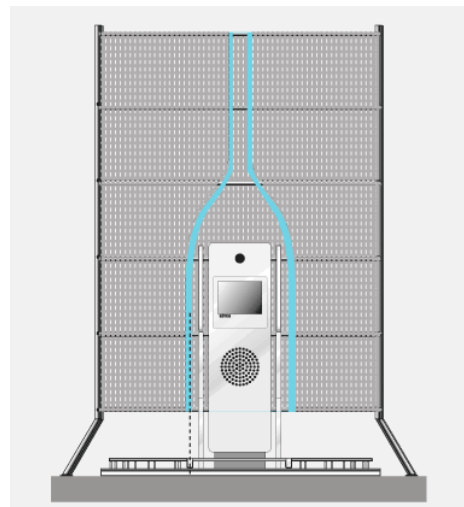


Figure 4 : atMOS Design

In this system, an imaginary XY level is structured, by an image processing technique, on the image taken by the digital video camera. The XY surface level is applied to the LED pointer, which keeps moving in the XY surface and is being held by the user. Based on these coordinates, the amount of user's movement is detected by calculating the distance (\square) in a certain sampling span. This enhances the reproducibility of body movements.

LED pointers that emits a certain amount of light, which makes the image analyzing detect an accurate data, plays a very important role in the reproducibility of body sensation. At the same time, LED pointers can intensively add colors to the movie source taken by the digital camera. This enhances the expressivity.

The digitalized data of the user's action is sent to the real-time effect section and is given to a parameter for the sounds and images to be processed.

There are also 8 pressure sensors (Figure 5) arranged in a circular pattern in the floor of the chassis. Each one of these sensors is allocated to a set of image effect and sound effect. Each sensor has a different shaped sticker, so that the user can remember which place to step on when she wants to use the effect. This enhances not only the reproducibility of the intention but also the reproducibility of body sensation in this content.

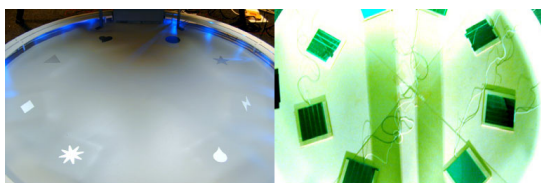


Figure 5 : Pressure Sensor

3.2.2 Real-Time Effect Section

The real-time effect section processes the sound and image using the data of the amount of movement as formed in the performance section.

In the real-time effect section, the objective was to create a system with high expressivity which flows together with the music, so that users can choreograph himself/herself radically by the users' action and by selecting the movie effects and imaginary instruments.

Figure 6 is a sample of a movie made by atMOS. In this passage, the movie effects will be described. The movie processing for atMOS is done by a movie processing sub-system "Viscore". The movie material of the user taken by the digital video camera facing the user inside the chassis is processed in real-time.

8 separate effects are prepared for each pressure sensors in the floor of the chassis. A complex effect can be processed by a combination of these 8 effects, and this enhances the expressivity. As the users are able to experience an arbitrary enjoyment, the expressivity of intention is enhanced.

In addition to this, the amount of movement obtained in the performance section is connected with the intensity of the effect, so the reproducibility of intention is enhanced by this collaboration. The intensity of the effect will rise when the user moves intensely.

The 8 movie effects have a common meaning for each row to enhance the reproducibility of body sensation. By making the effects have a common meaning for each row, the user can re-experience the same situation by perceiving which sensor has the intended effect, and this can be checked on the monitor at real-time. The front row sensors are effects to change the hue of colors. The sensors in the second row are those to process multiplication effects. This can be identified clearly on the monitor while the user is playing at a short distance from the monitor. The sensors in the third row are effects with afterimages. This effect can be enjoyed by viewing the monitor from the rear of the chassis. The sensors in the fourth row are effects to change the movie by dividing it into a number of pieces. The user stands at the rear end of the chassis and will be able to enjoy the shapely change of image in the monitor.

In this passage, the processing method for sounds will be mentioned. The sound processing for atMOS is done by a sound processing sub-system "Toad&Spider". This system is an automatic music creating engine which automatically



Figure 6 : Vizcore Image Sample

processes score creation, real-time synthesis and real-time effects. This tries to automatically complete all the process needed to compose trance/techno music. Toad&Spider is starts simultaneously with the initialization of atMOS. The expressivity is enhanced as a different music is created each time.

Considering the completeness of the musical piece, the interactivity was designed simple and to the minimum. The user controls 8 interactive imaginary instruments. These are allocated on each of the pressure sensors in the floor. As the effects are allocated on the same sensors as the movie effects, so these effects are grouped in a similar concept. This enables the high reproducibility of intention with the imaginary instruments by keeping the speed, chords at a certain level and by changing the scale of the music.

Also, atMOS has an imaginary audience system. The cheering sound is given as a effect to the users' movement. More the user intensively moves in the chassis, more the imaginary audience applauds.

3.2.3 Rendering Section

The rendering section encodes the movie data sent from the performance section to 3GPP, a standard format for 3G cellular phones. 3GPP is adopted to enhance the exchangeability, because this enables the movie to be seen not only with cellular phones but also with PC and other digital devices. Figure 7 shows an example of how the movie, created by atMOS, is played on a cellular phone.



Figure 7 : Movie on 3G Cellular Phone

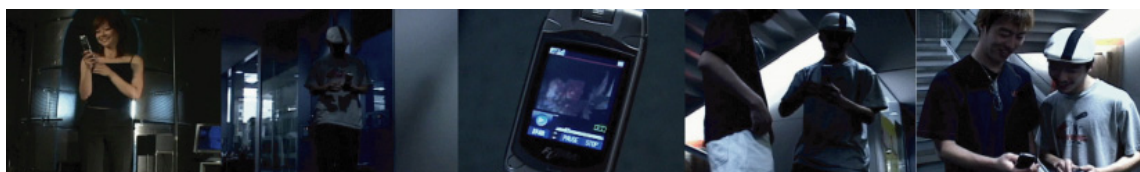


Figure 8 : Communication Image

3.2.4 Communication Section

The communication section uploads the encoded movie to the server and sends a mail with the URL for download to the users' cellular phone. Users is able to enjoy the movie not only by himself/herself but also with other people by transferring it by mail as seen in Figure 8. Likewise, users can enjoy other users' movie created by atMOS, by downloading it to one's cellular phone. atMOS enhances the exchangeability by evolving the playing, download, viewing and sending in a seamless flow, just as image communication evolves seamlessly from shooting to sending.

3.3 Verification

This chapter verifies the effectiveness of atMOS as a self expression movie generating system and an encouragement to mobile movie communication through the data obtained by the two exhibits.

3.3.1 SIGGRAPH 2003

During the days of exhibition at SIGGRAPH, 300 people played atMOS. The ratio of men to women who played atMOS was 1:1.

As this exhibition itself was based on new technologies, people who came to this exhibition did not seem to feel reluctant to move around in the chassis. Without explaining the details for atMOS, people who actively moved their bodies while playing noticed the movie to have a strong difference. This proves that the change of effects encourages the user in a sense of reproducibility of intention.

A large number of sample data was obtained from this exhibition. There was a great difference between each movies of the users. This supports that the realization of expressivity and reproducibility was achieved in each section. This realization makes clear that there was originality and rarity in each of the user's content.

As this exhibition was in USA, the movies could not be sent to the user's cellular phone by network. Instead, the data was copied to a 3G cellular phone, which the users viewed, through a memory card. Despite this fact, the movie was very popular and many users asked for the movie to be sent to their PC. Also, some people who had already played atMOS appeared again and brought their friends to the booth. As the system was assumed to be played by a number of people at the same time, by being able to process multiple LED pointers movement, there were many cases where users played this with their friends. This verifies the effectiveness of encouraging direct communication.

Figure 9 is an example of the movie with especially high originality. Owing that people did not feel reluctant to dancing, there were many movies that had highly original effects synchronized to user's original movement.

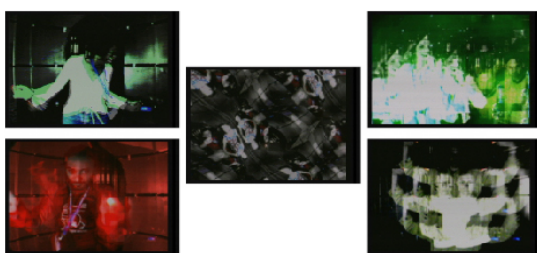


Figure 9 : Self Packaging Movie Sample

3.3.2 Japan Media Art Festival

In this exhibition, the complete system was presented as the system was fully equipped with network and cellular phones. Given this factor, we did a questionnaire survey to verify the effectiveness of mobile communication media. Figure 10 is a graph from the result of the survey. The users' intent of whether the user wants the movie, if the user wants to show the movie to others, and if the user wants to see other people's movie. These questions were answered in 4 levels (Absolutely

Yes, Yes, No, Absolutely No) and 132 people of ages 13 to 40 answered this question.

72% of the users answered that they want their movies. 56% of the users answered that they want to show this movie to others. And 79% of the users answered that they want to see other peoples movies. As a result, a high number of users wants their own or other peoples' movies but only about half the users wants to show it to others. Based on this result, the effectiveness of this system as a mobile movie communication can be verified that though only half of the users wish to exchange their movies, they feel a satisfaction to the mobile movie created by this system, and if the movie resulted as a cool movie the users feel that they would exchange it with others. Based on this standpoint, the aim of this research was accomplished by the product which is created by this system.

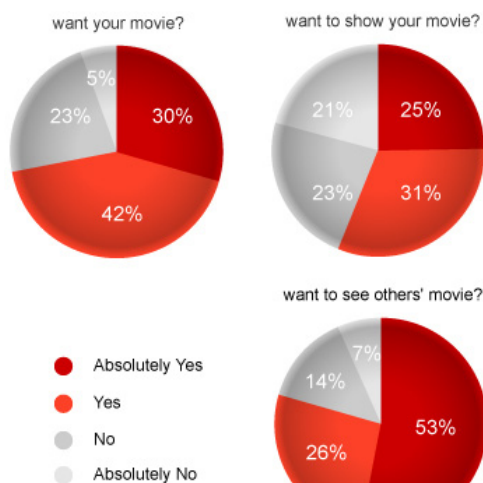


Figure 10 : Graph for User's Intent

On the other hand, the attribution of the source of this questionnaire survey must be taken into account. Although this exhibition was an art festival, and most people who came there were interested in digital contents, there was a different aspect from the exhibition in USA. That is to say, there were only a few people who would play this system actively as seen at the exhibition in USA. Due to the difference in national character, there is a need to verify the effectiveness by researching in other countries of Asia, Europe and USA as a comparative study.

3.4 Observation

Due to the attainment of reproducibility, expressivity in each section, atMOS enables to make a mobile movie full of originality. The rarity and originality of exchangeability was realized due to the high expressivity based on the reproduction of the users' experience on atMOS.

The system, on the other hand, still needs improvement. In order to enhance the exchangeability, an enhancement of reproducibility and expressivity, which forms the originality and rarity, is necessary. For example, in the present system, the reproduction of motive action is remained only to reproduce the motion of the hand holding the LED pointer, but there is a need to approach from what the movement implies. For instance, the movie undulate a wave pattern when a user moves one's arm in a wave like motion.

Also the exchange should not be concluded by an analog method of mail, and there is a need to set a place of exchange outside the system. For example, a device that enables to obtain other peoples' movies by accessing to a certain community site can be considered. It is essential that there is a need to structure a new method of communication that involves many people, which was not possible when the tool for communication was by mail that was sent mainly to acquaintances.

4 Conclusion

This research focused on movies as a new communication tool, and proposed a self expression movie generating system "atMOS". The purpose of this research was to encourage mobile movie communication and adopted the concept of exchangeability, expressivity and reproducibility for mobile movie contents in the system. This successfully demonstrated the potency of mobile movie communication.

Reference

- 1.<http://www.casio.co.jp/k-tai/a5406ca/detail1.html>
- 2.http://www.nttdocomo.co.jp/english/p_s/products/index.html
- 3.http://foma.nttdocomo.co.jp/english/tv/list/tv_tv.html
- 4.<http://www.vodafone.jp/english/products/kisyu/v601n/index.html>
- 5.[http://www.nttdocomo.com/presscenter/pressreleases/press/pressrelease.html?param\[no\]=457](http://www.nttdocomo.com/presscenter/pressreleases/press/pressrelease.html?param[no]=457)
- 6.http://www.kddi.com/english/corporate/news_release/2003/1006/attachment.html
- 7.<http://www.au.kddi.com/win/ezweb/ezchannel/index.html>
- 8.Robert St. Amant, Thomas E. Horton, and Frank E. Ritter. 2004. Model-Based Evaluation of Cell Phone Menu Interaction. *Proceedings of the 2004 conference on Human factors in computing systems*, 343 – 350.
- 9.Andriy Pavlovych, and Wolfgang Stuerzlinger. 2004. Model for non-Expert Text Entry Speed on 12-Button, *Proceedings of the 2004 conference on Human factors in computing systems*, 351-358.
- 10.Takeshi Kubo. 2003. "MIVE": A Proposal of LIVE Cinema Watching System using the Internet. *Keio University*.
- 11.Haruhiko Katayose, Tsuyoshi Miyamichi, and Nakamura Mitsuda. 2003. Collaborative Visual Jockey using Mobile Phone. *HCI International 2003*.
- 12.Ueda, N., Nakanishi, Y., Manabe, R., Motoe, M., and Matsukawa, S. 2003. GIS with cellular phone + WebGIS - Construction of WebGIS using the GPS camera cellular phone. *The Institute of Electronics, Information and Communication Engineers*.
- 13.Yoshimasa Niwa, Takafumi Iwai, Yuichiro Haraguchi, and Masa Inakage. 2004. The Re: living Map - an effective experience with GPS tracking and photographs. *Pervasive2004*.
- 14.James E. Katz, and Mark A. Aakhus.2002. Perpetual Contact: Mobile Communication, Private Talk, Public Performance. *Cambridge Univ Pr*.
- 15.<http://www.atlus.co.jp/am/english.html>
16. http://www.vodafone.jp/english/live/shamail_s/index.html
- 17.http://www.nokia.com/nokia/0_43613.00.html
- 18.<http://www.siggraph.org/s2003/conference/etech/atmos.html>
- 19.http://plaza.bunka.go.jp/english/festival/sakuhin_backnumber/15/atmos.html