weather-vox central: Promotion of Mood Sharing by Extraction of Mood Information from Voice and Presentation to a Shared Space

Sho Sakurai[1], Takatsugu Kuriyama[2], Takuji Narumi[1], Yasuo Kuniyoshi[2], Masato Sasaki[3] and Michitaka Hirose[1]

The University of Tokyo

ABSTRACT

We made a system named "weather-vox central" as an interface to support understanding others' mood. Besides a conversation, this system extracts mood information from vocal information. Then, it presents extracted mood information to the talking place with a common expression that is easily interpreted. This offers a common measure for speculate others' mood and promotes the intuitive understanding and sharing others' mood. Here, we use the looks of the sky as a representation of mood information. Since, the look of the sky is compared to a mood state. Moreover, it could influence mood. The users compared own mood state with the other's one, and talked about the reason of the mood, speculating each other's mood from the look of the sky without asking each other of the mood directly. In this way, a common measure of the personal mood was provided by this system. The system successfully supported users to share the mood.

KEYWORDS: Extracting Mood Information, Visualization, Implicit Information, Mood Sharing.

INDEX TERMS: B.4.1 [INPUT/OUTPUT AND DATA COMMUNICATIONS]: Data Communications Devises—Transmitter; J.5. [Computer Application]: ARTS AND HUMANITIES—Arts

1 INTRODUCTION

Mood, affect or atmosphere is information that individuals recognize subjectively. When persons try to share these information, it is necessary to translate them into some kind of symbols and concrete description[1]. However, it is not easy to translate into language and share these information, especially about mood, because mood is subjective, vague, and nonverbal information. Mood is personal experience reflecting past experience, and a measure to speculate mood state differs among people[2]. In addition, mood is feeble feeling and changes gently. It tends to be dependent on a vague expression, for example, "woolly-minded" or "athrill", and these are equivocal. This is a point different from emotion. Emotion is temporary strong feelings, so that it is expressed and categorized, for instance, as "happy" or "sad"[3]. Furthermore, mood is affected by environment and changes unconsciously. Therefore there are many cases that he/she cannot explain own mood state well.

The authors aim to support sharing such mood information by techniques to extract them automatically. In late years, many sensing technologies for various environmental information and systems to present its information have been proposed. For example, The BikeNet mobile sensing system which captures a cyclist performance to by various sensors and a system to sense[4] and visualize CO2[5]. In this regard, however, what these technologies and systems deal with is the information which can be captured with physical measures. Put simply, this information is easily converted into numerical number. Thus, previously proposed systems cannot be applied to deal subjective information, since it cannot be described with numerical value. However, we can feel information like CO_2 concentration or temperature even though these are numerical data. we can feel other's mood state in own bones from some sort of information intuitively, such as facial expression or vocal sound so that we feel CO_2 concentration or temperature from breathlessness or wind-chill factor. From here onwards, we consider that there are ways to sense mood information. Then, we think that a problem is that numerical number or linguistic expressions are unfitted as a way to transmit subjective information to others.

From this, we form following hypothesis. First, it is possible to extract mood information and measure to speculate mood state to many people. For this, it is required to capture features that vary in accordance with mood state variation. Second, the mood information is represented in a way where many people can easily interpret, and the promotes intuitive share of mood information. Moreover, it can present non-explicit mood information explicitly.

In this article, we made a system named "weather-vox central". This is a system as an interface mounts two functions that it can achieve above-mentioned two hypotheses to inspect the above-mentioned hypothesis. Then, we observe reactions of people who experienced "weather-vox central".

2 RELATED WORK

Many studies analyze vocal information do not depend on remark contents as the non-language information to measure his/her emotion[6][7][8]. "kokoro-scan", which commercialized evaluates the various items of emotion[9]. Besides, this gives a person reading a result a standard to judge his/her state of emotions or atmosphere. This system evaluates the various items of feelings. Moreover, this gives a person reading a result a standard to judge his/her state of feelings or atmosphere. However, what these study concern chiefly is not vague mood but categorized emotions. Furthermore, the information shown is inorganic expression such as numerical value or graph. It is dubious that these expressions are able to express sensuous information intuitively.

On the other hand, as concern as studies of interfaces for sensing and sharing of subjective information, for instance, Shuzo's group makes "Shoji"[10]. This system analyzes one's mood or environmental atmosphere that is analyzed from vital information or temperature and presents there with some color or tremor. Likewise, Tsujita et al. proposed "Syncdecor" system[11]. This transmits others' ambient information or presence sensibility

by motion of livingwares. However, these systems are not considered to measure mood state because these aim to transmit ambient information of a person to others in a remote place.

In this article, we propose a system which presents mood information by a expression which is associated with mood. This led many people to understand intuitively and share other's mood.

3 EXTRACTIING MOOD INFORMATION BASED ON VOCAL INFORMATION AND ADOPTING AN EXPRESSION TO PROMOTE HAVING COMMON INTERPRETATION OF PEOPLE

3.1 System Summary of weather-vox central

The authors made "weather-vox central"(fig.1). This is a system estimates mood state by analyzing voice volume and pitch in conversation. In addition, this presents information of mood state by change of look of the sky on a screen.

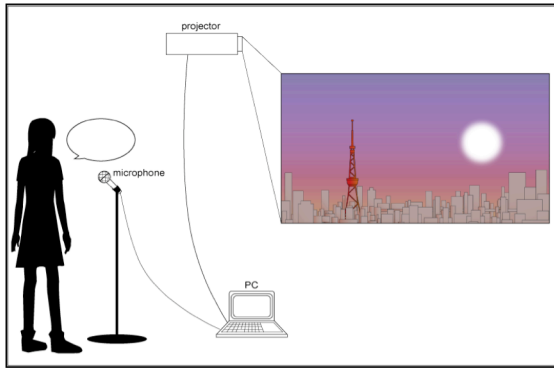This system comprises a PC, a microphone and a projector.

Figure 1. System of weather-vox central

3.2 A Method to Extract Mood Information

"weather-vox central" extracts mood information automatically. This "mood" is change of affects on a axis of pleasant/unpleasant, based on a circumplex model of affect that was advocated J. A. Russell[12]. The concrete method to extract mood information is as follows.

Many people tend to speak in clearly and briskly with improving intonation when they feel pleasant like a fun or happiness. On the contrary, they tend to talk in a dark and monotone voice when they feel unpleasant like anxious or depression. For this, the system assumes the speaker's mood as pleasant state when he/she speaks in a loud or high voice, with intonation. Conversely, his/her mood is assumed as unpleasant state when he/she speaks in a deep, low or monotone voice.

Firstly, the voice taken through a microphone is FFT analyzed and we set four indexes: Volume of the voice, pitch of the voice, the volume and the pitch in every 1/60 second. These indexes enable to measure the volume and the pitch of the voice, and intonation. In addition, from these indexes, we set mood-value. This value is a parameter to represent the change of his/her mood. This value is set 50 as a neutral state. The neutral case that his/her mood judged mostly unchanged. The value is normalized within 0 to 100, and 50 is for neutral. Then mood-value is lower than 50, his/her mood is speculated as unpleasant state. Besides, mood-value is higher than 50, his/her mood is assumed as pleasant state.

Moreover, mood-value are calculated as following. 0-860[Hz] part of the input voice is divided into twenty 43[Hz]-width bands. This maximum value, 860[Hz] is the upper limit value which the frequency where reaction is seen in conspicuously when some people speak. Then, define values of the amplitude of the sound of

each band are a_1, a_2, …, a_{20}. Then, define a value of the entire amplitude as the volume of the voice, A. A is calculated as (1). In addition, define values of the amplitude of the sound in each band as f_1, f_2, …, f_{20}. F, that is defined as average pitch of the voice is calculated as (2).

$$A = \sum_{i=1}^{20} a_i \qquad (1)$$

$$F = \sum_{i=1}^{20} \frac{f_i a_i}{A} \qquad (2)$$

Furthermore, define values of the volume calculated every 1/60 second sequentially as A_1, A_2, …, A_t. Then, an enhanced absolute value of the fluctuation of A is Ad. Ad_t is calculated as (3). Similarly, define values of pitch of voice calculated every 1/60 second sequentially as F_1, F_2, …, F_t. Then, an enhanced absolute value of the fluctuation of F is Fd_t . Fd_t is calculated as (4).

$$Ad_t = \left\{ A_t - A_{t-1} \right\}^2 \qquad (3)$$

$$Fd_t = \left\{ F_t - F_{t-1} \right\}^2 \qquad (4)$$

Here, define mood-value that is calculated every 1/60 second is M_t. Moreover, when mood value is 50, in other words, when a mood state is neutral, define $A_t=N_A$, $F_t=N_F$, $Ad=N_{Ad}$, $Fd=N_{Fd}$. At the time of $A_t<N_A$ and $Ad_t\leqq N_{Ad}$, a mood state is estimated to be a negative state. Then, in the case of mood state changing into a negative state, M_t is calculated after distribution in two patterns because two types of case are showed.

Here, we set these values followed the change of real mood state ont the basis of preliminary tests.

One is slow change, and the other is rapid change. Therefore, we divided a calculation method into two ways when changing to a negative state. If $F_t<N_F$ and $Fd_t\leqq N_{Fd}$, a mood state is estimated that changes to a negative state slowly, M_t is calculated as (5). Likewise, if $F_t\geqq N_F$ and $Fd_t>N_{Fd}$, a mood state is estimated that changes to negative state rapidly, M_t is calculated as (6).

$$M_t = M_{t-1} - \frac{A_t}{10} \qquad (5)$$

$$M_t = M_{t-1} - \frac{A_t}{15} \qquad (6)$$

Oppositely, when mood state changes into positive states, M_t is calculated without distribution because mood state changes rapidly in most case. When $A_t>N_A$ and $Ad_t>N_{Ad}$, a mood state is estimated that changes to positive state, M_t is calculated as (7).

$$M_t = M_{t-1} + \frac{A_t}{20} \qquad (7)$$

3.3 Look of the Sky as Expression to promote Mood Sharing

We adopted the change of the look of the sky in a city as expression to express the change of mood state intuitively. This reason is so that it is thought that the sky state seems to affects mood or emotion to many people. For example, on a fine day and the clear day of the air, he/she feels better. On the contrary, his/her mood is low on a day weather is bad or night. Further, because the sky are often used as a metaphor of the mood state. Such examples that mood is likened to weather has been supposed in much poetry and literary works[13].

In reference to the above, sky graphic is showed that is change with changing mood state. namely the mood-value by this system.

In reference to the above, sky graphic is showed that is change with changing mood state. namely the mood-value by this system. The gently changing of the mood is expressed visually and some people can interpret in the same way for a personal mood state.
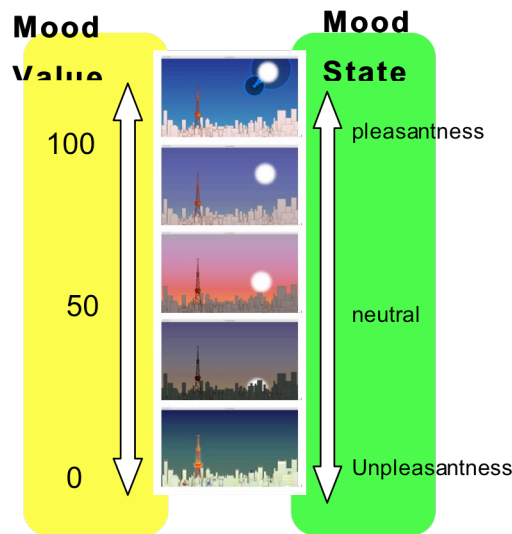

Figure 2. Relation of Mood-value and change of look of the sky in graphic

4 PRESENTATION OF MOOD INFORMATION TO SPACE MORE THAN ONE PERSON

4.1 Experimental Exhibition

The authors exhibit 'weather-vox central' at an art exhibition 'iii Exhibition 9' hold in The University of Tokyo in July 2008 (fig.3)[14]. Then, we observed reactions of users, in other words, speaker and people who shared space with speaker, and communication between them, such as conversation.

4.2 Reaction of People who Experienced weather-vox central

In above-mentioned Exhibition, about 70% persons who spoke at front of "weather-vox central" made remarks that the system reflected own mood state in their conversations. Specifically, they remarked that the position of the Sun, color of sky or speed of changing of them represent their pleasantness/unpleasantness. In addition, conversations as premises for feedback of changing of the graphic were revolved. For example, when sky in graphic was sunny when a speaker speaks, people with the speaker asked the reason why he/she is pleasant to him/her. Then, the speaker

Figure 3. Pictures of iii Exhibition 9

answer he/she considers to be the reason of own mood state, for instance, "I'll go to Disneyland tomorrow." or "I had a date with my boyfriend." On the contrary, sky in graphic change to night, people with the speaker question a subject about the reason of him/her unpleasantness. Then the speaker answered like such follow, "I have had much work to do." or "I have an exam." Thus, people shared space with the speaker ask the reason why the speaker is pleasant/unpleasant after they view the feedback of graphic. At this time, there is no scene where speakers react negatively toward the feedback of graphic.

Here, in this 70% of people who experienced this system, there were speakers operated a tone of voice intentionally. As one of example, night came in graphic when one person said he is looking forward to tomorrow's work in a clear and firm voice consciously. At this time, he answered tomorrow's work makes me feel gloomy in actuality. Moreover, two persons with the speaker asked him why he was unpleasant after they saw the feedback of graphic even though expected the feedback of graphic from his vocal sound was not shown. The two persons said they could not guess original mood state of the speaker.

Additionally, scenes were observed that some persons compared own mood state to others' mood state based on condition or speed of changing of the sky and position in the Sun. They were talking about the reason why he/she is more pleasant/unpleasant than oneself. In this time, the speaker and people with the speaker said that look of the sky is expression as a way to present mood state of individuals comprehensibly.

However, there were people who experienced weather-vox central who said this system did not reflect own mood state. As concern as these patterns, night falls though some persons were good mood. Or, the Sun in screen rose even though some people were blue mood. In these cases, some people said because he/she has negative thinking or he/she has talked in more a low voice than others.

Others, when a laughter or cheer has scattered, mood-value reached a maximum and the Sun in screen rose the highest position, regardless of how to speak or tone of voice of the speaker. Additionally, there were cases that a person who is not an acquaintance bat shared space with a speaker addressed and asked about him/her mood state after the person saw the feedback of the

sky graphic to the speaker. Thus, new communications were created via this system.

5 DISCUSSION

About 70% users remark that the reaction of sky graphic synchronized with his/her mood state of own. From this, the method of analysis of mood state based on a tendency to utterance propose in this article is can become the one of effective method for sensing mood information. Furthermore, there were remarks that the sky graphic that reflected above-mentioned analysis results are easy to understand and conversation between users assuming the reaction of the graphic. It is thought that these suggest that this system enables to give a standard to recognize the difference of mood of users. On top of that, users had Common interpretation for shown information and shared mood information between plural users.

However, the extraction method of mood information we proposed in this article is not considered about the individual quality of voice. Therefore it is thought that some users of this system were not analyzed definitely.

6 CONCLUSION

In this article, we proposed a method of extraction mood information and presentation it by using graphic of the look of the sky to express intuitively. These propose aimed sensing mood information automatically and supporting to sharing it between multiple persons. We made a system named "weather-vox central". This system mount functions to carry out a method we proposed and presents mood information to a shared space. Through the observation, it is suggested that a method to extraction and presentation mood information we proposed is effective to describe mood state comprehensible and promote sharing it. However, we need to consider quality of voice of individual.

We can catch sensuously others' mood or atmosphere others are feeling without vocal information. In future, we want to examine the reason and investigate methods to create environments that many people can share mood or other subjective information.

REFERENCES

[1] C. E. Shannon and W. Weaver. The Mathematical Theory of Communication. *The University of Illinois Press*, Urbana, Illinois, 1949.

[2] R. S. Lazarus and S. Folkman. Stress, Appraisal, and Coping. New York, Springer, 1984.

[3] J. P. Forgas. Affect in social judgment and decisions: A multiprocess model. In: M. Zanna (ed), *Advances in experimental social psychology*. vol.25, Academic Press, San Diego, pages. 227–275, 1992.

[4] S. B. Eisenman, E. Miluzzo, N. D. LANE, R. A. Peterson, AHN, G-S, and A. Campbell. The BikeNet mobile sensing system for cyclist experience mapping. *Proceedings of the 5th international conference on Embedded net-worked sensor systems*, pages. 87-101, 2007.

[5] T Nariya, YA Seong, T Hashida and T. Naemura. Spatio-temporal sensing and visualizing of CO2. *International Conference on Computer Graphics and Interactive Techniques SIGGRAPH '09 Posters,* New Orleans, Louisiana, Article No.95, 2009.

[6] T. Johnstone and K. R. Scherer. The effects of emotions on voice quality. *Proceedings of the 14th International Conference of Phonetic Sciences*, pages. 2029–2032. San Francisco, 1999.

[7] R. Nakatsu, J. Nicholson and N, Tosa. Emotion recognition and its application to computer agents with spontaneous interactive capabilities. *Knowledge-Based Systems*. volume.13, No.7-8, pp.497-504, 2000.

[8] K. R. Scherer. Vocal communication of emotion: A review of research paradigms. Speech Communication, volume.40, No.1-2, pages 227-256. April 2003.

[9] http://kokoro.sega.jp/ (accessed 2010-07-25)

[10] M. Shuzo, M. Shimura, Y. Motohashi, J. J. Delaunay, I. Yamada. A Communication Terminal for Sending and Receiving Ambient Information. Korea-Japan Design Engineering Workshop 2008, Seoul, July 9-10, 2008.

[11] H. Tsujita, K Tsukada, I Siio. SyncdDecor: Communication Appliances for Couples Separated by Distance. UBICOMM 2008, *The 2nd International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, pp.279-286, 2008.

[12] J. A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology* 39, pages.1161–1178, 1980.

[13] R. Proudfoot, A. Thompson and D. S. Kasten (ed.). The Arden Shakespeare Complete Works. Revised Edition. Thomson Learning, London, 2001.

[14] http://i3e.iii.u-tokyo.ac.jp/i3e9/ (accessed 2010-07-15)