

Development of Tool for Video-Debriefing for Understanding Emotion over Usability

Kun-Pyo Lee*, Sang-Hoon Jeong**

* Department of Industrial Design, Korea Advanced Institute of Science and Technology

Address: 373-1, Kusong-dong, Yuseong-gu, Daejeon, 305-701, Korea

Tel: +82-42-869-4514 / Fax: +82-42-869-4510 / E-mail: kplee@kaist.ac.kr

** Department of Industrial Design, College of Fine Arts, MOKWON University

Address: 800, Doan-dong, Seo-gu, Daejeon, 302-729, Korea

Tel: +82-42-829-7947 / Fax: +82-42-822-7950 / E-mail: diasoul@mokwon.ac.kr

- Type of paper: Full paper for oral presentation
- Key conference theme: Design and emotion: methodological issues

Abstract

In general, the methods for measuring human emotion are divided in two categories, one is the psychological way based on user's subjective evaluation and the other is the physiological way based on physiological signals. However, these methods have some well known limitations. In this research we suggested the methods for measuring user's emotion in the natural and accessible environment for the design field. In order to observe user's emotional changes while they interact with a product, the video-debriefing and emotion logging program "VideoTAME" was developed as an effective measurement tool for complementing the current psychological methods. In the Testing Module of VideoTAME, subjects engage in 'Video-Debriefing' after usability testing: They evaluate their emotional changes through playing and watching the video clips of their performing tasks in the experiment room. In the Analyzing Module, the researchers replay the results created by subjects during the experiment and analyze the results using Microsoft Excel. It is hoped that this study will be of great help for designers for effective measurement of user's emotions naturally expressed while using a product.

Key words: tool for measuring user's emotions, usability, user's emotions

1. Introduction

Human emotion is the issue which is subjective, difficult to define and even more difficult to measure because it is personal. Cacioppo and Gardner (1999) concluded that “the measurement of emotion is a bustling research area.” Generally, the methods for measuring human emotion are divided in two, one is the psychological way based on user's subjective evaluation and another is the physiological way based on physiological signals.

Traditionally, the measurement of emotion is usually dependent on the psychological method (Plutchik, 2003). The most commonly used psychological method requires users to report their emotions with the use of a set of adjective interval scales like semantic differential methods or verbal protocols. Various researches have been done on psychological measuring methods in areas such as psychology, linguistics, emotional engineering, etc. However, this type of method, based on subjective evaluation, has some limitations. When studying what users say about their emotional state, user's emotions tend to change according to time. Plus, precisely expressing emotions in words is not an easy task. And, because emotional experiences can be delicate, users sometimes (consciously or unconsciously) manipulate their reports. Most of all, since the subjective report is examined after the emotional part is experienced, the emotion expressed can be different before and after the examination. Rosenberg and Ekman (1994) showed that these kinds of self-report do get influenced from recent effects. According to Scherer (1986), self-reports are subject to distortions due to ego-defence tendencies and socially desirable effects. Moreover, relying on the subject's memory is also a significant limitation. There is always a time-lapse between the experienced emotion and the self-report. The bigger the time-lapse between experience and report, the more the results are distorted.

The physiological method, measuring physiological signals to use them as an objective index of human emotion, has been tried to measure the emotion more objectively. However, the method using the physiological signals also has problems. First, it is very burdensome and unnatural with recent technology. For example, a subject has to have sensors on his fingers to monitor the pulse and the electrodermal activity (EDA), and also has to have equipments on his head to measure an electroencephalogram (EEG) coming from the occipital and parietal lobe. Secondly, there are limitations in the stimulus to make subjects express their emotion. The stimulus is limited to one that makes subjects not need to move, such as visual or hearing equipments. Especially, to measure an EEG, the control of other waves has to be severe. It is still difficult to detect a pure EEG generated from emotion because eye blinking or tiny movements can make considerable noise. The third problem is

an economic issue. High-costing equipments are required to sensor physiological signals and there are a lot of troublesome matters to furnish the testing environment. Fourth problem is the lack of consistency in the results of emotion research by measuring physiological signals. Even though it has consistent results about electrocardiogram (ECG) or EDA, it still requires an objective index about other signals. Since the stimulus and the physiological variables are different by researches, the research results by measuring methods are inconsistent (Collet et al., 1997). The problems on these research methodologies are the difficulties of setting the framework for the emotion-specific reaction pattern and the relationship between emotion and autonomic nervous system reaction (Prkachin et al., 1999). Therefore, it can be said that it is not easy to understand human emotion just by physiological signals. Lastly, there are also difficulties in data analysis. It is not easy to approach the area from the design field since it requires propound knowledge of human physiological signals to analyze data from the tests.

Therefore, it is necessary to research on the methods for measuring user's emotion in the natural and accessible environment for the design field and for measuring user's emotion naturally expressed while using a product, and to solve problems of the two existing methods. The aim of this research is to suggest a most effective method for measuring user's emotion expressed while interacting with a product. If taking into consideration that the stimulus is not big, and the user has to keep moving while using the product, to measure the user's emotions, the psychological measuring method can be seen as more effective than the physiological measuring method. In our previous research, we extracted a set of emotional words and representative emotions that can show changes in emotions of users while using a product and not words that were derived simply from the product's appearance (Jeong & Lee, 2005). In this research, we have come up with a suitable set of subjective evaluation scale for measuring user's emotions expressed while using a product. Through this, we suggest a more effective emotion measuring method that can complement the limitations of the psychological measuring method. The developed software named VideoTAME (Video Think Aloud for Measuring Emotion), basically asks the subject to watch a recorded video clip of the user performing a specific task and to examine the emotional changes that had occurred.

2. Objectives of development and structure of tool

To solve the problems mentioned in the existing methods for measuring emotions, the following objectives for development were derived to effectively measure user's emotions:

- Since emotion has to be measured after the emotion has been felt and experienced, emotion distortion should be minimized.
- The environment where the product is used should be as realistic as possible.
- User's emotion should be able to be measured in a natural environment where product is used.
- Environment should be easily accessible for the design field and should be able to sufficiently gather enough opinions from subjects.
- Collected data should be able to be transferred and utilized to other commonly used programs.
- Collected results should be databased and able to be taken out again to be used effectively.
- Suitable emotional words should be suggested to suit the various words users might come up with while using product.

The video-debriefing tool for understanding over usability was developed taking the objectives listed above into consideration. The tool called VideoTAME is composed largely of a Testing Module and an Analyzing Module (See Figure 1). The Testing Module is composed of one part where the subject actually performs the experiment and the other part where the data is collected and recorded. The Analyzing Module is composed of a database part which can classify, store, and manage each experiment, and an analyzing part that plays and analyzes the collected data in various ways.

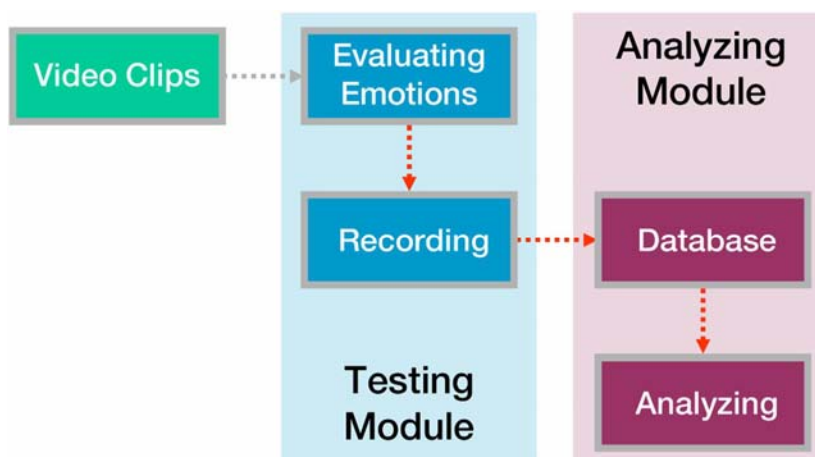


Figure 1: Structure of user emotion evaluation tool

VideoTAME was developed using Microsoft Visual-Basic 6.0. A distinct characteristic is that it comes with 2 Windows Media Player engines. It is also compatible with various video clip formats such as AVI, MPG, WMV, MOV, etc. When the subject ends the session, the details of the experiment is recorded and saved in a text file format (*.txt). The saved file is then reworked as database in the Analyzing Module. The biggest characteristics of the Analyzing Module are that specific scenes can be captured from the video clip, and that all sorts of analyzed data can be directly converted into Microsoft Excel files by Excel Object 9.0. In addition to this, the subject can stop the video clip and then the selected emotional words automatically go into the sound folder. The designer designs it so that when it plays, the emotional words can come out as voice recording so that when the subject performs the task it can give a “think aloud” effect. The following is a rough overview of the development environment of VideoTAME:

- Development Language: Microsoft Visual-Basic 6.0
- Used Engine: Microsoft Windows Media Player 9.0, Excel Object 9.0
- Development Environment: CPU Intel® Pentium® 4 CPU 2.80GHz, 512MB, XP Service Pack2
- Operating Environment: CPU Intel® Pentium® 2 CPU 600Mhz, 128MB, WIN98 and above
- Size of Program: Testing Module (228KB), Analyzing Module (1.25MB)
- Program Structure: Testing Module (Video TAME.exe), Analyzing Module (VideoTAME-Analyzer.exe), guide.wav, Sound Folder

3. Testing Module of VideoTAME

This is the Testing Module where the subject performs the experiment by following the guiding message. The Testing Module starts as the subject selects the folder where the recorded file of the subject performing the task in the experiment room is located. The Testing Module is composed of four screens. An initial screen that shows basic information on general performance of experiment, a screen that shows the movie clip that the subject can see and evaluate emotional changes, a screen used after the experiment for emotion evaluation using the subjective evaluation scale and a screen with a brief survey.

In the part where the emotion is tested, the experiment subject examines his/her own emotional change by watching the video clip that was recorded in the experiment room when the task was performed.

The video clip is composed of a part where the subject's facial expression can be seen and a part where the actual product can be seen. First of all, a task that can examine emotion should be selected, and then while playing the video clip, when there is visible emotional change the clip should be paused, then the subject should select all the emotions he/she felt from the list of the 32 emotional words. Color is added to the selected emotional words so that it can be easily distinguished (See Figure 2).



Figure 2: Emotion evaluation screen and selected emotional words

Multiple answers can be selected for the emotional words, and after selected, subjects are guided to express how strong they felt in the selected emotional words by 5-point scale. The items of the scale are automatically changed according to the emotional word selected. At the end, the subject gets to briefly write down why emotional change had occurred (See Figure 3).

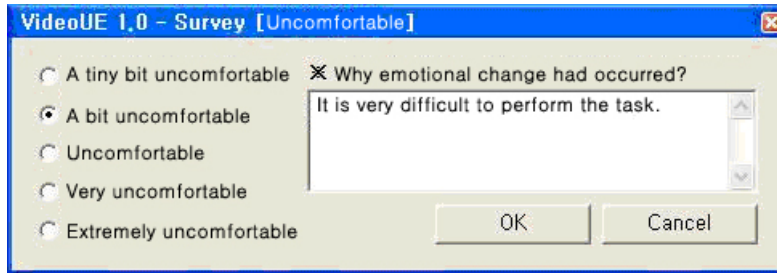


Figure 3: Screen showing emotional state and reason for selecting the specific state

Additional words can be selected while the clip is paused, and if there are none, the video clip can be resumed and the emotion evaluation test can be proceeded in the same method as above. Evaluations can be added for previous screens by rewinding the video clip, and revisions can be made for the previously selected emotional words. If all emotion evaluation of the performed task is completed, the screen goes to the next screen. In the survey that is shown at the end of the experiment, a general evaluation is given about the product applied by the subjective evaluation scale made in the previous our research (See Figure 4).

Finally, the experiment ends with some basic questions related to the subject (gender, age, academic background, occupation) and some questions related to the used product. For example, if the experiment was to evaluate emotion expressed while using a cellular phone, the following questions could be asked.

- How long have you been using a cellular phone?
- What previous types of cellular phone have you used? Please list all the company names.
- What type of cellular phone are you currently using? Please write down company name, model type, and purchased date.

VideoTAME 1.0 - 주관적 평가스케일

The following are emotional words that can express a human's psychological state or feeling while using a product. Please check on all the emotions you felt while performing the task (more than one answer is allowed) and select one that you think fits the best among the 5-point scale. (1. A tiny bit true, 2. A bit true, 3. True, 4. Very true, 5. Extremely true).

No.	Emotional word	Evaluation	No.	Emotional word	Evaluation
1	Pretty	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	17	Elegant	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
2	Satisfactory	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	18	Natural	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
3	Unusual	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	19	Annoyed	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
4	Discomfort	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	20	Glad	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
5	Joyful	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	21	Clean	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
6	Excellent	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	22	Reliable	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
7	Careful	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	23	Unique	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
8	Smart	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	24	Lovable	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
9	Uneasy	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	25	Efficient	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
10	Pleasure	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	26	Marvelous	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
11	The best	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	27	Confused	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
12	Refined	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	28	Valuable	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
13	Original	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	29	Disappointed	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
14	Difficult	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	30	Cute	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
15	Like	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	31	Lovable	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
16	Look better	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	32	Harmonious	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5

OK

Figure 4: Subjective evaluation scale

4. Analyzing Module of VideoTAME

The Analyzing Module is where the designer or the researcher can play or analyze the collected data in various ways. The Analyzing Module starts as the designer selects the folder that contains the evaluation results of the Testing Module (See Figure 5). The Analyzing Module is composed of two screens: A screen that allows the designer to examine the experiment results by playing the recorded video clip by each task, and a screen that variously analyzes the results that are exported to Microsoft Excel.

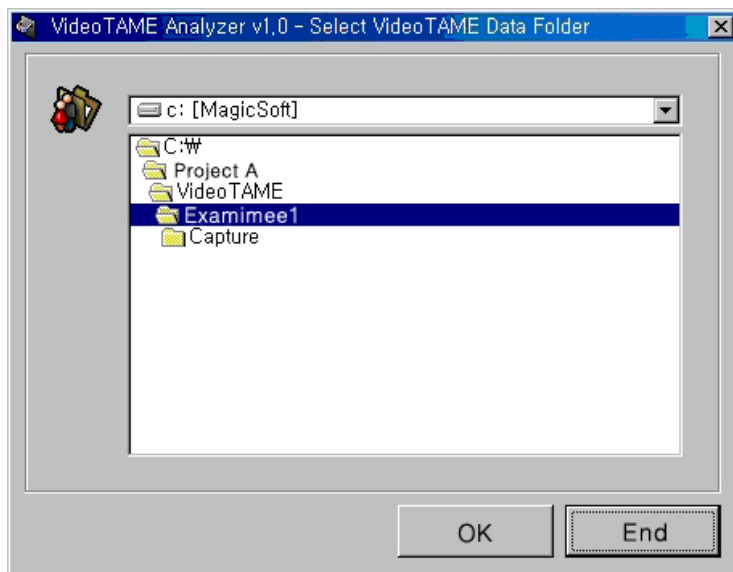


Figure 5: Screen of selecting folder to be analyzed

In the screen where the experiment results can be seen by each task, the designer can check the evaluated emotional change (See Figure 6). The designer can easily recognize the part where the subject selected an emotional word by stopping the video clip because it automatically indicates it. Plus, the selected emotional words and the degree of it gets displayed at the bottom of the screen and lets the designer know through voice recording so that the designer can have a “think aloud” effect of the performed experiment. Other than this, the selected emotional words, the score given for the selected degree, the reason for choosing the specific word, and the average of the representative emotions can be identified through the selected emotional words. Lastly, when the analyzer finds something particular while playing the video clip, the situation can be input by clicking the “Description of Situation” button. The information that was input can be re-checked through the exported Microsoft Excel file.



Figure 6: Experiment results per task

Through the exported Microsoft Excel file, variously analyzed information can be seen as follows.

•**Analyzed information per scene:** User's emotional change per situation can be analyzed by the following – the subject's task division, time that video clip was stopped to select emotional word, scene, information on explaining situation, emotional words and degree selected from subjects, reason for selecting those specific words, and the average of the representative emotions can be identified through the selected emotional words (See Figure 7).




Microsoft Excel - Age28_Male_8 years_060411093602								
L11								
Examinee Information								
Gender: Male Age: 28 Academic Background: Graduate Student Occupation: Graduate Student Using Period: 8 years								
Previous Phone: LG, SKY, Motorola Current Phone: Motorola, MS-150, August, 2003								
Exported Time: 2006-04-11 AM 9:35:15								
Task	Time	Scene	Situation Description	Selected Words (Score)	Total	Reason for Selecting	Representative Emotion	Average
1	00:00:27		The user is now using the main menu scene.		0		Aesthetics	0
				Satisfactory(2)	2	It is similar to the interface of the current phone I am using.	Satisfaction in usability	2
					0		Novelty	0
				Difficult(2)	2	It is very difficult to perform the task.	Uncomfortableness	2
					0		Pleasure	0
					0		Excellence	0
1	00:00:52				0		Aesthetics	0
					0		Satisfaction in usability	0
					0		Novelty	0

Figure 7: Analyzing screen per scene

•**Analyzed Information per task:** User's emotions for each task can be analyzed by checking the average score of emotional words, number of times of emotional words selected, and the average score of representative emotions. The information on analysis per task applies not only to the subject but also to the whole group of subjects (See Figure 8).

Representative Emotion	Emotional Word	Number of Times Emotional Word Selected	Average	Average of Representative Emotion
Aesthetics	pretty	0	0	0
	refined	0	0	
	elegant	0	0	
	clean	0	0	
	lovable	0	0	
	cute	0	0	
	sweet	0	0	
harmonious	0	0		
Satisfaction in usability	satisfactory	2	2	2
	careful	0	0	
	natural	0	0	
	reliable	0	0	

Figure 8: Analysis screen per task (each subject)

•**Information analyzing the whole experiment:** The emotional change occurred from the user using the product can be analyzed through the whole task, the number of times emotional words selected, it's average score and the average of the representative emotions can be identified through the selected emotional words. After the experiment, the results while using the product and after using the product can be compared showing the average scores of representative emotions and emotional words selected from the subjective evaluation scale. For analysis of the entire experiment, not only can the subject be analyzed but the whole group of subjects can be analyzed too (See Figure 9).

Representative Emotion		Emotional Word	Number of Times Emotional Word Selected	Average	Average of Representative Emotion	Subjective Evaluation Scale	
						Average	Average of Representative Emotion
Aesthetics		pretty	0	0	0	1	1.8
		refined	0	0		2	
		elegant	0	0		0	
		clean	0	0		3	
		lovable	0	0		0	
		cute	0	0		2	
		sweet	0	0		0	
		harmonious	0	0		1	
Satisfaction in usability		satisfactory	10	2.7	2.75	2	2.5
		careful	0	0		2	
		natural	5	3.2		2	
		reliable	1	3		3	
		efficient	3	2.33333333		3	
	valuable	0	0		3		

Figure 9: Analysis of whole experiment (all subjects)

•**Analysis of subject's basic information:** Information on subject's serial number, gender, age, academic background, occupation, etc. and answers to questions related to the experiment is displayed. These data can be used in various ways to analyze the results of experiment through cross-tabulation analysis.

5. Conclusion

In this study, we have suggested methods for measuring a user's emotions in the natural and accessible environment of the design field. First, by extracting emotional words and representative emotions that show changes of user's emotions while using a product, we came up with a subjective evaluation scale for measuring the user's emotion. Through this we developed a tool that effectively measures the user's emotions expressed while using a product and that can complement the limitations of the psychological measuring method. In the Testing Module of the developed software, the subject can view the recorded video clip of he/she performing some specific task, and examine the emotional changes occurred. In the Analyzing Module, results can be verified by running the total experiment clip by each task, and then the data can be sent to Microsoft Excel and be analyzed in various ways. The tool developed through this research, can be used to effectively measure a user's naturally expressed emotion while using a product. To summarize the significance of the development of this tool would be as follows:

- Firstly, a more realistic product-using environment is suggested by letting the subject to examine his/her own emotional changes occurred while performing the task by actually watching the video clip that was recorded. Distortion on emotional measurement that Scherer (1986) had once pointed out as the limitations of psychological measurement can also be minimized.
- Second, the emotions naturally expressed by users in a natural environment can be evaluated effectively by complementing the previously burdensome and unnatural equipments needed for the physiological measuring method.
- Thirdly, by easily analyzing and collecting various data on facial expressions and reasons on why the emotion was felt, it has become easier to be approached from the design field.
- Fourth, by comparing the situation while using the product and the emotional words selected in that situation, it suggests the possibility of analyzing the emotional change occurred from user, and the usability of the product.

Through this research we expect to clarify what effect a product's usability has on user's emotions with the help of a tool developed for measuring user's emotions while using a product.

References

Cacioppo, J. T. And Gardner, W. L. (1999) "Emotion." Annual Review of Psychology 50: 191-214.

Plutchik, R. (2003) *Emotions and Life: Perspectives From Psychology, Biology, and Evolution*. Washington: American Psychological Association.

Rosenberg, E. L. And Ekman, P. (1994) "Coherence between expressive and experimental systems in emotion." Cognition and Emotion 8: 201-229.

Scherer, K. R. (1986), "Studying emotion empirically: issues and a paradigm for research." In Scherer, K. R., Wallbott, H. G., and Summerfield, A. B. (eds) *Experiencing emotion: a cross-cultural study*. Cambridge: Cambridge University Press, 3-27.

Collet, C., Vernet-Maury, E., Delhomme, G., and Dittmar, A. (1997) "Autonomic nervous system response patterns specificity to basic emotions." Journal of the autonomic nervous system 62(1/2): 45-57.

Prkachin, K.M., Williams-Avery, R. M., Zwaal, C., and Mills D. E. (1999) "Cardiovascular changes during induced emotion: an application of lang's theory of emotional imagery." Journal of Psychosomatic Research 47(3): 255-267.

Jeong, S. H. And Lee, K. P. (2005) "Extraction of user's representative emotions expressed while using a product." Journal of Korean Society of Design Science 18(1): 69-80.