# International Journal of Human Computer Interaction (IJHCI)

ISSN : 2180-1347

Volume 1, Issue 3

Number of issues per year: 6

# International Journal of Human Computer Interaction (IJHCI)

## Volume 1, Issue 3, 2010

**Edited By**
**Computer Science Journals**
**www.cscjournals.org**

# International Journal of Human Computer Interaction (IJHCI)

Book: 2010 Volume 1, Issue 3
Publishing Date: 20-12-2010
Proceedings
ISSN (Online): 2180-1347



IJHCI Journal is a part of CSC Publishers
http://www.cscjournals.org


Published in Malaysia

Typesetting: Camera-ready by author, data conversation by CSC Publishing Services – CSC Journals, Malaysia

**CSC Publishers**

# Editorial Preface

This is third issue of volume one of International Journal of Human Computer Interaction (IJHCI). IJHCI is an International refereed journal for publication of current research in Human Computer Interaction. Publications of IJHCI are beneficial for researchers, academics, scholars, advanced students, practitioners, and those seeking an update on current experience, state of the art research theories and future prospects in relation to applied science. Some important topics covers by IJHCI are affective computing, agent models co-ordination and communication, computer mediated communication, innovative interaction techniques and user interface prototyping for interactive systems etc.

This journal publishes new dissertations and state of the art research to target its readership that not only includes researchers, industrialists and scientist but also advanced students and practitioners. IJHCI seeks to promote and disseminate knowledge in the applied sciences, natural and social sciences industrial research materials science and technology, energy technology and society including impacts on the environment, climate, security, and economy, environmental sciences, physics of the games, creativity and new product development, professional ethics, hydrology and water resources, wind energy.

IJHCI editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the Journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc, Scribd, CiteSeerX and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJHCI. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. IJHCI provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

**Editorial Board Members**
International Journal of Human Computer Interaction (IJHCI)

# Editorial Board

# Table of Content

Volume 1, Issue 3, December 2010

Pages

# Toward a More Robust Usability concept with Perceived Enjoyment in the context of mobile multimedia service

**Jieun Sung** sungji1@msu.edu
*Telecommunication, Information Studies & Media*
*Michigan State University*
*East Lansing, USA*

**Younghwa Yun** younghwa@msu.edu
*Telecommunication, Information Studies & Media*
*Michigan State University*
*East Lansing, USA*

**Abstract**

Mobile multimedia service is relatively new but has quickly dominated people's lives, especially among young people. To explain this popularity, this study applies and modifies the Technology Acceptance Model (TAM) to propose a research model and conduct an empirical study. The goal of study is to examine the role of Perceived Enjoyment (PE) and what determinants can contribute to PE in the context of using mobile multimedia service. The result indicates that PE is influencing on Perceived Usefulness (PU) and Perceived Ease of Use (PEOU) and directly Behavior Intention (BI). Aesthetics and flow are key determinants to explain Perceived Enjoyment (PE) in mobile multimedia usage.

**Keywords:** TAM, Perceived Enjoyment, Mobile multimedia, Aesthetics, Flow, Social Presence.

## 1. INTRODUCTION

The traditional usability concepts have been dominated by instrumentally motivated ideas of information technology use. Approaches to measure usability have mainly focused on the tasks, the goals, and their efficient accomplishment. However, since new technology such as the invasion of mobile phones into people's daily lives, the complexity of products is increasing and user's expectation becomes challenge with using new technology. According to Manovich [1], interactions with computers and computer-based devices penetrate people's lives outside of work. Especially, the mobile phone is getting attractive to customers with its aesthetic features such as animated icons and sounds, the personalized interfaces, and the various shapes and surface finishes. Also, a mobile phone came to be used for all kinds of non-work activities: entertainment (game, music, video and TV), information searching and social life, because of its multi-functionality and flexibility of its usage. As a result, the emphasis on efficiency and functionality came to be replaced by new criteria such as being friendly, pleasurable, aesthetically pleasing, animated graphics [1]. With this viewpoint, narrow focus on task-related usability was challenged to designers and developers and widened by introducing "emotional usability" [2]. In the emotional usability concept, enjoyment is more related with an important determinant of the adoption of mobile service than the usefulness [3, 4, 5]. This study attempts to investigate aesthetic, flow, and social presence as significant predictors of usability in multimedia use of mobile context, especially in the aspect of Perceived Enjoyment (PE). The model integrates Technology Acceptance Model [6], social presence theory [7], theory of flow [8], and aesthetic. By investigating the linkage between these three concepts and PE, this study examines considerable predictors for more robust usability of mobile phone usage.

## 2. PROBLEMS IN TRADITIONAL USABILITY CONCEPT

Traditionally, usability is defined as being:

"…the effectiveness, efficiency and satisfaction with which specified users achieve specified goals in particular environment." [9]

Effectiveness refers to the extent to which a goal or a task is achieved. Efficiency is about the amount of effort required to accomplish a goal. Satisfaction refers to the level of comfort that a user feels when using a product and to how acceptable a product is to users for achieving their goal [10]. Thus, evaluation of usability has been conducted by having an evaluator to record the data such as time per task, the number and types of errors and user attitudes while users are doing some tasks. Likewise, Nielsen mentioned [11], [12] that usability was the question of how well users could use the functionality and the main goal of the website should make users complete their tasks easily and fast even if there was a need for fun and time to enjoy or relax on the web. As shown above, there are rooms to argue the concept of traditional usability for information technology use. Approaches to measure usability have mainly focused on given tasks and goals, and their efficient and effective accomplishment.

### 2.1 Effectiveness

In a case of task oriented usability, scoring of effectiveness is arguable in some contexts such as creative production, management, data analysis. Also, it is not clear in the traditional usability concept of how to measure what outcomes are effective for users and, how users and outcomes are related [13]. Also, as new technologies have embedded into the daily lives, the complexity of product is increasing and user's expectations become higher. For example, mobile phone can be simultaneously used while travelling, talking or doing other activities. However, most usability tests assume that the context of a user facing a product would come with the full attention in comfortable environments. In the various contexts, the meaningful outcomes from users could be different from the effectiveness of classic usability [14].

### 2.2 Efficiency

Efficiency is evaluated by the invested time or effort during completing the given tasks. In other words, interactions between the products and users are evaluated by the time, number of steps, or number of variations from ideal path. Those variables are highly correlated but there are exceptions. For example, there can be a step users spend most of time or just skip not because of the efficiency, but because of past experiences or preferences such as fun or special meaning to users [13]. Usability issues are increasingly demanding and complex and are better considered of as being about consumer experience than ease of use. The new technologies and applications now have more challenge on users' expectations, various types of users, and usages. However, existing usability concepts are unable to handle such an intricate and multifaceted definition. For example, IPTV (Internet Protocol TV) or DiTV (Digital interactive TV) is not just a TV. They attempt to provide interactions between TV and viewers. Existing usability cannot measure how users can best manage those interactions because it just emphasize on the functionality of TV [14]. Also, mobile phone gives challenge to the traditional concept of usability. In addition to the context, existing usability cannot explain why people are so immersed into texting even though numeric keyboard is not effective and efficient as a view of usability. There are needs for understanding users' behaviors based upon user experience integrated with, and beyond functionality [15].

### 2.3 Satisfaction

The term "satisfaction" in usability is concerned with evading negative feelings rather than producing positive emotions. It indicates that good usability equals the removing of usability flaws

[16]. The concept of "satisfaction" is not enough to cover the human factors for usability. Also, many new technologies are associated with both work and leisure, not just with task-based performance issues. Based upon the user experience with performance, the satisfaction is influenced by preferred user experience, aesthetics of product, intrinsic value to the users [13].

## 3. Extended Technology Acceptance Model (TAM) for the mobile multimedia

### 3.1 Mobile Multimedia Services

Even only a decade ago, all mobile services were simple communication-oriented and network-based applications, such as voice calls or SMS messages. However, people have been accustomed by the content-oriented services and various multimedia solutions. 3G, recently including some 4G networks and smart phones with various value-added services are introduced and reaching mass mobile markets. It is widely understood that the emergence of multimedia services has reflective implications in the mobile industry based on the rise of IP-based mobile services [27]. The mobile multimedia services refer to a trend in which mobile services integrate with improving the contextual value through entertainment, information or communication [28, 29, 30]. In this study, all services and applications dealing with graphics, music, game, and video are included in the scope. Also communication oriented functions which relate to multimedia content are in the scope of the study. These include;

- Music/audio players (both offline and streaming)
- Movie/video players (both offline and streaming)
- Photo/image viewers
- Photo functions (Phone camera)
- Blogging applications
- Searching information applications (clients providing access to online information)
- Games
- Multimedia communication (MMS, Bluetooth, IM, email)

### 3.2 Extended Technology Acceptance Model

Traditional Technology Acceptance Model (TAM) suggests that Perceived Usefulness (PU) and Perceived Ease of Use (PEOU) of Information Technology (IT) are major determinants of its usage. Davis [6] defined PU as "the degree of which a person believes that using a particular system would enhance his or her job performance," and PEOU as "the degree of which a person believes that using a particular system would be free of effort." Also, user's beliefs or trusts determine the attitude toward actual using the system. Behavioral intentions (BI) to use are determined by these attitudes toward using the system. Finally, BI to use leads to actual system use. With this traditional TAM, PEOU and PU constructs have been considered important criteria in determining the acceptance and use of IT in the past decades (Keil et al., 1995; Malhotra & Galletta, 1999; Moon & Kim, 2001). From the same aspect, the traditional approaches toward improving usability have been focused on ease of use. The Human-Computer Interaction (HCI) research has focused primarily more on technology-centered aspects of usability, such as time complete tasks and the number of errors. These variables are the most fundamental and motivational factors consisting of the TAM. Information system researchers have investigated and agreed that PEOU and PU are valid in predicting the adoption of various information technologies [19, 20, 21, 22, 23, 24].

However, depending on the specific technology context, additional explanatory variables may be needed beyond the ease of use and usefulness constructs. Davis [6] argued that future technology acceptance research needs to address how other variables affect usefulness, ease of use, and user acceptance. Factors which are antecedent of the acceptance of a new IT are vary with the technology characteristics, target users, and context [19]. Especially, mobile phones provide pleasure and usefulness at the same time. In other words, users will expect to get information and enjoyment at anytime and anywhere.

According to Vorderer et al. [25], Perceived Enjoyment (PE) is the core of media entertainment experience, and it can be found in many ways depending on the user's readiness and ability to suspend disbelief to engage. PE has been confirmed that it has an important role in user technology acceptance and has great implications [6]. Also, PE is based on the user experience such as "how the person felt about the experience, what it meant to them, whether it was important to them, and whether it sat comfortably with their other values and goals." [26]. This study would add PE as a significant predictor of behavior intention, and suggest aesthetic, flow, and social presence as additional factors of explaining PE in the mobile multimedia usage. The proposed research model of Extended TAM for this study is shown in Figure 1.

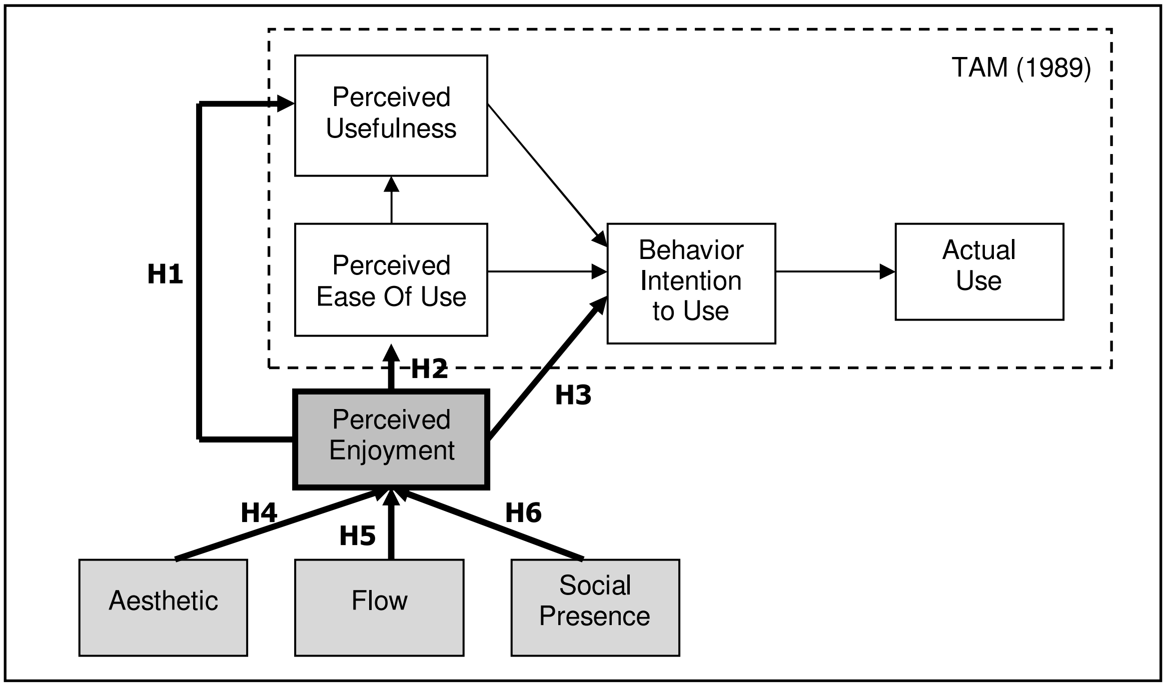



**FIGURE 1:** Proposed research model of Extended TAM for the mobile multimedia services.

### 3.3 Perceived Enjoyment (PE) → Perceived Usefulness (PU)

PE is defined as "the extent to which the activity of using computers is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated" [31]. PE is referred to as an intrinsic motivation variable such as the doing of an activity for satisfactions rather than for some outcomes or results. In contrast, extrinsic motivation, such as PU is a construct that measures how user's productivity and effectiveness have been improved by using the product [32]. Davis et al [31] found that usefulness and enjoyment were significant determinants of behavioral intention. However, the effect of enjoyment on perceived usefulness was relatively unexamined [33].

People with a pleasurable perception of the enjoyment from using the product are more likely to perceive it useful [8, 34]. Also, Agarwal and Karahanna [35] found a multi-dimensional construct called "cognitive absorption," a state of involvement with software, had a significant influence on PU. High cognitive absorption status which makes high impact on PU is enjoyment. Assuming other things being equal, the more enjoyable a product is the more useful a product can be perceived. Furthermore, the purpose of mobile multimedia service includes pleasure and enjoyment such as music, photo, movie, or even game. It means that PE would affect positively on PU in mobile multimedia services. To verify the facts, following hypotheses is addressed.

H1: PE has a significant positive effect on PU in the context of mobile multimedia services.

### 3.4 Perceived Enjoyment (PE) → Perceived Ease Of Use (PEOU)

PEOU is defined as “the degree to which a person believes that using a particular system would be free of effort” [6]. Also, PE has been conceptualized as an antecedent of PEOU (PE→PEOU) [33, 36, 37]. Previous literature revealed that the causal direction between PE and PEOU had been proposed and confirmed (Table 1). Studies using this direction usually referred to the technology acceptance model (TAM) with the justification that PE makes users underestimate the difficulty with using the technologies, because they enjoy the process itself and do not perceive it to be difficult [36]. Also, positive affective emotion makes users perceive themselves as having generous time to complete a task and it actually reduces the perception of workload related with using the technologies [33, 35]. Also, Huang et al. [38] examined that PE has a positive impact on PEOU in the context of mobile learning. Thus, the study proposes the following hypothesis:

H2: PE has a significant positive effect on PEOU in the context of mobile multimedia services.

| **PE → PEOU** | | | |
|---|---|---|---|
| **Article** | **Model** | **Context** | **Findings** |
| [35] | Extended TAM | Web | PE is one dimension of Cognitive Absorption (CA); CA has a significant impact on PEOU. |
| [36] | TAM | Online help desk, multimedia system for property management | PE influenced PU via PEOU, without assessing its direct effect on PU over and above PEOU.<br>PE’s impacts on PEOU increase along with increasing experience. |
| [37] | TAM & motivational model | Virtual Workplace | PE does not have direct impacts on BI. Instead, its effects are fully mediated by PU and PEOU. |
| [33] | Extended TAM | Web-based class management system | PE has significant effects on PEOU and PU. |
| [45] | Expectation-Confirmation Model (ECM) | IT product/ service | Users’ PE of IT is positively related to their PEOU of IT. |
| [102] | TAM | Virtual work place | PE has a significant impact PEOU in game-based training. |

**TABLE 1:** Review of the existing literature on the causal relationship between PE and PEOU.

### 3.5 Perceived Enjoyment (PE) → Behavior Intention (BI)

Enjoyment is significant not only in offline settings such as shopping [39, 40], but also in online contexts [41]. Prior research suggested that enjoyment directly affected the BI of online customers [42]. For example, in the research of online consumer behavior, Koufaris [43] found that shopping enjoyment played a critical role in predicting consumer intention to return to an online store. Also, Li et al. [44] found that users who perceived the use of IM (Instant Messaging) as enjoyable were more likely to intend to continue using it. For the IT product and services, users’ PE had a significant effect on user’s intention to use [45]. In the online gaming context, one important motive for playing online games was to seek pleasure or enjoyment; players who experienced enjoyment and the emotional response of pleasure were more likely to be motivated to play more [46, 47, 48]. Prior research also showed that enjoyment could indirectly impact BI through other variables. For example, Venkatesh [36] examined factors of PEOU and found that PE significantly impacted on BI to use information technology through PEOU. Accordingly, the following hypothesis is addressed:

H3: PE affects BI in the context of mobile multimedia services

### 3.6 Aesthetic → Perceived Enjoyment (PE)

Interface design is increasingly important as both information and entertainment websites in order to compete for rapidly increased and complicated customers [49]. The sensory experience of the website can also determine whether a user stays, shops or revisits [50, 51]. Especially, online retail shopping sites have suggested to include both utilitarian and hedonic dimensions and vendors can create aesthetically rich shopping environments to provide enjoyment to customers. Like the examples of online shipping environment, the aesthetics is considered to make the positive emotions and lead a more favorable judgment towards a system or technical products in HCI research [52, 53]. Also, according to Postrel [54], aesthetics has intrinsic value above and beyond their functional values, and it satisfies users' needs and motivates them. Users want to achieve not only certain well-defined goals, but also involve the sense of affective responses [55, 56, 57].

Heijden [58] proposed a new concept, Perceived Attractiveness," for the TAM model in the Internet context. It was defined as ''the degree to which a person believes that the website is aesthetically pleasing to the eye.'' He found empirical support that perceived attractiveness of the website did influence PU, PE, and PEOU [58]. Cyr et al. [3] also found that design aesthetics had a significant impact on PU, PEOU, and PE in the mobile commerce context. Also, the design aesthetics may have a larger relative impact on PE than on PU and PEOU. In this study, we draw on research on visual aesthetics and apply this work in the specific context of the mobile multimedia services. We expect that perceived visual aesthetics of the mobile interface would impact PE. Therefore, the following hypothesis is addressed.

H4: Aesthetic affects PE in the context of mobile multimedia services.

### 3.7 Flow → Perceived Enjoyment (PE)

Flow is a concept originally suggested in the 1970s [59, 60, 61] to explain the pleasure found by immersion in everyday activities. According to Csikszentmihalyi [62], the enjoyment was realized by artists when they were immersed in the creative act. Also, Hsu & Lu [63] found when people were in flow, they shifted into a common mode of experience and they became absorbed in their activities. In other words, it is felt when "...instead of being buffeted by anonymous forces, we do feel in control of our own fate. ...we feel a sense of exhilaration, a deep sense of enjoyment." In order to reach this state of optimal experience: "There must be a goal in a symbolic domain; there have to be rules, a goal, and a way of obtaining feedback." Users must be able to concentrate and interact with the opportunities at a level corresponding with their skills [8, 64].
Csikszentmihalyi [65] outlined nine dimensions of flow:

- Clear goals
- Immediate feedback
- Personal skills meet with challenges
- Merger of action and awareness
- Concentration on the task
- Control
- A loss of self-consciousness
- An altered sense of time
- Experience which becomes autotelic (self-contained goal experience)

Recently, flow has also been studied in the context of information technologies and recommended as a useful attribute in understanding consumer behavior [43, 66, 67, 68, 69, 70]. Also, some of the emotional and cognitive components used in flow research have values of intrinsic enjoyment, perceived control, and concentration/attention focus [71]. For many people, finding a moment of flow can be when they are doing things that they enjoy and excel at. In studies of games, Jones [16] reported that certain games made people so absorbed and concentrated that they could not stop playing. Also, Herz [72] found the intense captivation while people were playing a computer game. It is not just about play and competition. It is associated with the intense feelings of engagement that the game can stimulate and engender deep feelings the player. This type of devotion makes game to be perceived as a total package. It is not just the graphics, sounds, and other multimedia attributes. It is about how those attributes help, support, and give life to a domain that makes possible anything which has no counterpart in the physical world [64]. Like this example, flow theory has been applied to the game and well explained how the game captured the fun and entertainment motivation of users. According to Lu et al. [73], people were often in the state of flow when they used IM service. Also, when users log into an IM platform, they not only want to communicate with others, but also look for fun and try to obtain a flow experience. Therefore, the study hypothesizes:

H5: Flow positively affects PE in the context of mobile multimedia services.

### 3.8 Social Presence → Perceived Enjoyment (PE)

Social presence is defined as "being together with another" that a medium allows users to experience others like being psychologically present [74, 75]. Social presence theory considers social presence as an intrinsic component in a communication medium [7]. Some researchers emphasized on the psychological connection. They regarded social presence as concerned with "warmth." Encouraging a sense of human warmth and sociability can be achieved by providing actual interaction with other humans such as virtual communities or chats, or by instilling the imagination of interacting with other humans such as human audio and video, intelligent agent, or greeting with human touch [76]. Also, to stimulate the interaction with other humans through the medium, some researchers characterized the social presence of a medium as its capability to transmit information about non-verbal cues such as facial expressions and gesture [7]. Like this, both verbal and nonverbal cues contributed to social presence [77, 78], and the choice of communication medium was found impacting immediacy of social presence significantly [78, 79].
Recent studies [80, 81, 82, 83, 84, 85] addressed the importance of mobile telephone use in stimulating socially linked. The ability to communicate with anyone, anytime, and anywhere is the most fundamental value of telecommunication applications.

Mobile communication technologies provide easy ways of strengthening social bonds among friends, and even with strangers through functional and expressive management [85, 86]. The mobile communication has successfully moved people's communication patterns into a "social network- based paradigm," where people socially connected and interacted based on mutual trust, rather than being controlled by their physical location, mood or their appearance [85, 87]. For example, text-based IMs (Instant Message) allow users to express their emotional states by commenting their status or using emoticons which increase social presence. Also, people can use mobile phone anywhere, and they can find out other people's location status easily. More recently GPS (Global Positioning System) applications help to find nearby members of their social networks, leading to a physical meetings. Moreover, device status (on/off) might signal to others' availability and willingness to communicate [88]. Like this, the mobile system can enable users to integrate simulated social actors into their social network as easily as real social actors. As a result, mobile systems produce high social presence, and social presence mediates social interaction.

According to Biocca et al. [75], the most telecommunication bandwidth, such as internet and mobile, was used to gain satisfying, and productive access to others' thoughts, emotions, and presence of humans, because humans are social beings. Also, Heeter [89] found that users felt enjoyment more when they felt a stronger social presence in the context of a virtual reality entertainment system. With the same reason, the social presence was found consistently associated with enjoyment or fun [90]. However, there is little research about the relationship between social presence and enjoyment [90, 91]. Therefore, the study hypothesizes:

H6: Social presence positively affects PE in the context of mobile multimedia services.

# 4. RESEARCH METHOD

## 4.1 Respondents and Procedure

College students are an important population in studying mobile phone usage because they are not only familiar with cutting edge technology but also very excited to maximize mobile phones in daily lives. There were more than 262 million wireless mobile phone subscribers, representing 84 percent of the US population in June 2008, according to the Cellular Telecommunications and Internet Association [92], that is up 35 percent from 194 million in June 2005 and nearly three times more than the 97 million wireless subscribers in June 2000. In 2008, Harris Interactive reported that 9 out of 10 U.S. College Students own a mobile phone [93].

The M:Metrics report [94] found that employed college students are 42 percent more likely to use mobile email than the typical subscriber, 23 percent more than full-time workers. Working students download mobile games and personalize content on their phones twice as often as do other users. Many studies have shown that cell phone usage was subject to functional expansion, because users gradually changed habits and learned to apply the new technologies for a growing concerns and purposes in a wide range of situations [95].

A Web-based survey was employed for data collection. Undergraduate students at a large Mid-Western university participated in this study in exchange for extra credits. The survey has been conducted online (http://www.supersurvey.com) during nearly two weeks from November 18th to December, 1st. Total number of data collected was 148 and we used data from 129 respondents for this research after dropping useless data. However, among valid sample of 129, 16% of sample reported that they had a mobile phone without multimedia functions, yielding a total sample size of 108. The demographics indicated that there were more female (70%) than male (30%), consistent with the composition of the class. The average age of participants was 21 (*SD*=1.33).

## 4.2 Operational Measures

Measures were adopted from previous research and modified to fit the context of this research. All measurements are shown in Table 2. As for Flow construct, nine questions were adopted from previous related research [96]. In the original study, Chou & Ting [96] categorized flow into five groups; concentration, playfulness, time distortion, telepresence, and exploratory behavior. 1-2 items in each category have been selected for this study.

| Constructs | Measure Items *(7-point multi-item scales from Strongly disagree to Strongly agree)* | Sources |
|---|---|---|
| **Perceived Enjoyment (PE)** | ▪ I found using mobile multimedia services entertaining<br>▪ I found using mobile multimedia services pleasant<br>▪ I found using mobile multimedia services is exciting<br>▪ I found using mobile multimedia service is fun<br>▪ I found using mobile multimedia service is enjoyable. | [58, 97, 98] |
| **Perceived Ease of Use (PEOU)** | ▪ It will be impossible to use mobile multimedia services without expert help<br>▪ Learning to use mobile multimedia services is easy for me<br>▪ It is difficult to learn how to use mobile multimedia services<br>▪ I find it easy to use mobile multimedia services to do what I want it to do<br>▪ It takes too long a time to learn to use mobile multimedia services<br>▪ It is easy to remember how to use mobile multimedia services<br>▪ Using mobile multimedia services requires a lot of mental effort<br>▪ My interaction with mobile multimedia services is clear and understandable<br>▪ It is easy for me to become skilful at using mobile | [63, 97] |

| | | |
|---|---|---|
| | multimedia services | |
| **Perceived Usefulness (PU)** | ▪ Using mobile multimedia services improves my task quality<br>▪ Using mobile multimedia services improves the performance of my tasks<br>▪ Using mobile multimedia services supports the critical part of my tasks<br>▪ Using mobile multimedia services enables me to accomplish tasks more quickly<br>▪ Using mobile multimedia services increases my task productivity<br>▪ Using mobile multimedia services enables me to have more accurate information<br>▪ Using mobile multimedia services enables me to access a lot of information<br>▪ Using mobile multimedia services enables me to access the newest information<br>▪ Using mobile multimedia services enables me to acquire high quality Information | [63] |
| **Behavior Intention(BI)** | ▪ I will use mobile multimedia services on a regular basis in the future<br>▪ I will frequently use mobile multimedia services in the future<br>▪ I will use mobile multimedia services during the coming month<br>▪ I would spend money to mobile multimedia | [63, 99, 100] |
| **Aesthetic** | ▪ The screen design (i.e. colors, boxes, menus, etc) for mobile multimedia services is attractive<br>▪ The interface for mobile multimedia services looks professionally designed<br>▪ The graphics mobile multimedia services are meaningful<br>▪ The overall look and feel of the interface for mobile multimedia services is visually appealing | [3] |
| **Flow** | ▪ I am always totally fascinated with the game when using mobile multimedia services<br>▪ I am always absolutely focused on the game when using mobile multimedia services<br>▪ I experience the highest exciting when using mobile multimedia services<br>▪ I experience the highest satisfaction when using mobile multimedia services<br>▪ Time goes by very quickly when using mobile multimedia services<br>▪ I tend to forget where I am when using mobile multimedia services<br>▪ Mobile multimedia services provide me endless surprising behavior experiences<br>▪ I burned with curiosity when using mobile multimedia services<br>▪ Using mobile multimedia services make me feel like exploring a new world | [96] (selected) |
| **Social Presence** | ▪ There is a sense of human contact in the mobile multimedia services<br>▪ There is a sense of personalness in the mobile multimedia services<br>▪ There is a sense of sociability in the mobile multimedia services<br>▪ There is a sense of human warmth in the mobile multimedia services<br>▪ There is a sense of human sensitivity in the mobile multimedia services | [101] |

**TABLE 2:** Measure Items of main constructs

All items were scored on a 7-point Likert scale ranging from 1, strongly disagree, to 7, strongly agree. For measuring the mobile multimedia usage, respondents choose one of given answers from 1-3 times in a month to more than 4 times in a day and top three favorite features of mobile phone.

### 4.3 Data Analysis

PLS (Partial Least Squares) method is adopted to test the hypotheses and analyze data. PLS can provide not only the examinations of all paths in the proposed model (structure model), but also supplementary analyses with underlying items (measurement model). Éthier et al [103] mentioned that PLS is more prediction-oriented than other Structural Equation Model (SEM) tools. Also, PLS has the benefit with relatively small sample size [104]. Chin [104] recommends "rule of 10" guideline for PLS users: at least 10 cases per measured variable for the larger of (1) the largest latent factor block, or (2) the dependent variable with the largest number of incoming causal arrows in the model. In this research model, the second condition yields a minimum sample size required of 90, which is well exceeded by given sample size of 108.

The SmartPLS 2.0 software (http://www.smartpls.de) was used to analyze the measurement, structural model and SPSS17 was used to test construct validity and reliability.

## 5. RESULTS

26% of the respondents reported that they used mobile multimedia service 1-3 times in a month, while another 25% said that they used it more than 4 times in a day. There is a distinctly divided two groups to use mobile multimedia services. It is shown in Figure 2.

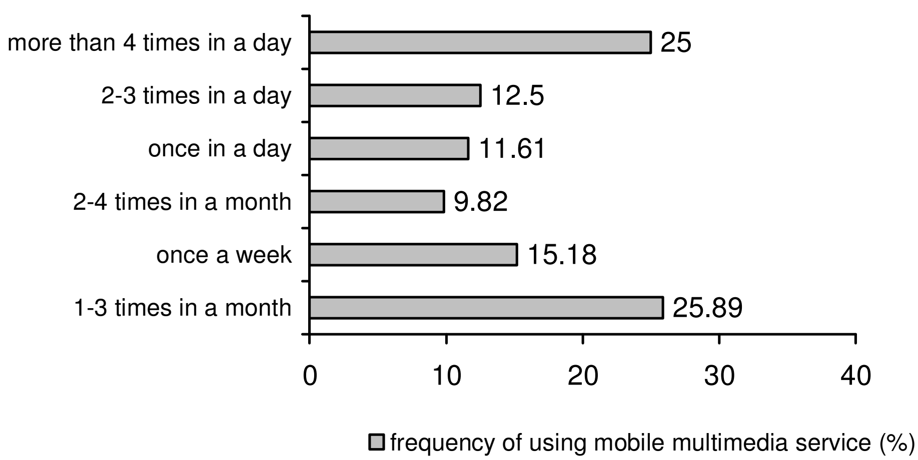



**FIGURE 2:** Frequency of using mobile multimedia service

The most popular mobile multimedia services were; phone camera (63.8%), photo image viewer (60.7%), MMS (Multimedia Messaging Service) and e-mails (52.7%) and music/audio players (50.9%). On the contrary, 3.6% of respondents reported to use blogging with their mobile phones (Figure 3).

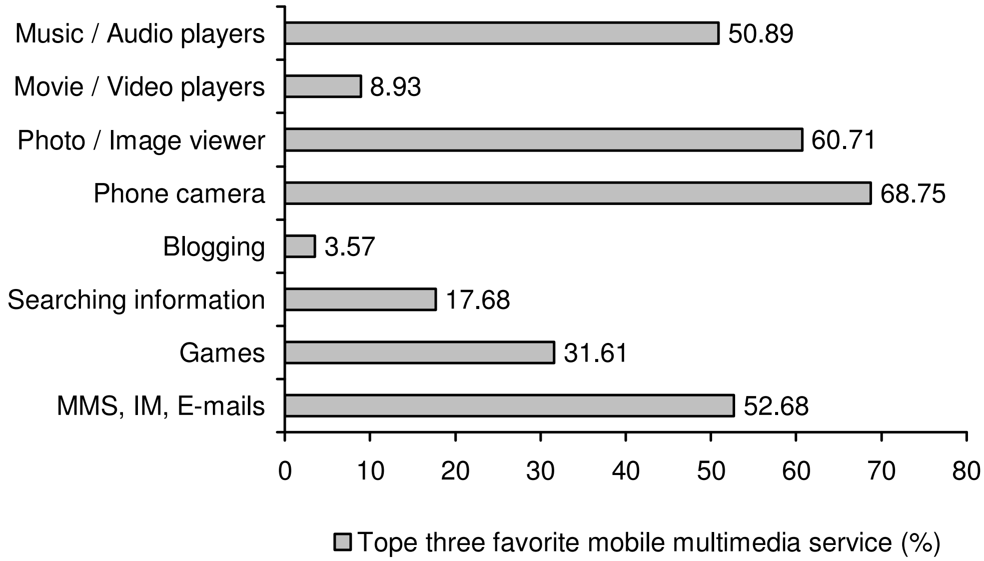



**FIGURE 3:** Top three favorite mobile multimedia services

Table 3 shows Crombach's α in each measurement. All composite reliability values range from .857 to .943, which were above .80. It shows that the construct reliability of our measurement model is acceptable.

| Measurement | N of items | Crombach α |
|---|---|---|
| Perceived Enjoyment | 5 | .943 |
| Perceived Ease Of Use | 9 | .907 |
| Perceived Usefulness | 9 | .929 |
| Behavior Intention | 4 | .857 |
| Aesthetic | 4 | .907 |
| Flow | 9 | .909 |
| Social Presence | 5 | .893 |

**TABLE3:** Crombach's α in Measurement

Table 4 shows a matrix of Pearson's Product-moment Correlation Coefficients. The relationships between Perceived Ease Of Use (PEOU) and Perceived Usefulness (PU) (p=.155), PEOU and flow (-.027) and SP (-.108) are not significant. Also there is no significant relationship between social presence and aesthetic (p=.157) at the .05 level of significance.

| | | PE | PEOU | PU | BI | AES | FLOW | SP |
|---|---|---|---|---|---|---|---|---|
| PE | Pearson Correlation | 1 | | | | | | |
| | Sig. (1-tailed) | | | | | | | |
| | N | 108 | | | | | | |
| PEOU | Pearson Correlation | .379** | 1 | | | | | |
| | Sig. (1-tailed) | .000 | | | | | | |
| | N | 108 | 108 | | | | | |
| PU | Pearson Correlation | .598** | .155 | 1 | | | | |
| | Sig. (1-tailed) | .000 | .055 | | | | | |
| | N | 108 | 108 | 108 | | | | |
| BI | Pearson Correlation | .555** | .376** | .573** | 1 | | | |
| | Sig. (1-tailed) | .000 | .000 | .000 | | | | |
| | N | 108 | 108 | 108 | 108 | | | |
| AES | Pearson Correlation | .344** | .447** | .352** | .364** | 1 | | |
| | Sig. (1-tailed) | .000 | .000 | .000 | .000 | | | |
| | N | 108 | 108 | 108 | 108 | 108 | | |
| FLOW | Pearson Correlation | .319** | -.027 | .442** | .362** | .308** | 1 | |
| | Sig. (1-tailed) | .000 | .392 | .000 | .000 | .001 | | |
| | N | 108 | 108 | 108 | 108 | 108 | 108 | |
| SP | Pearson Correlation | .212* | -.108 | .229** | .193* | .157 | .548** | 1 |
| | Sig. (1-tailed) | .014 | .134 | .009 | .022 | .052 | .000 | |
| | N | 108 | 108 | 108 | 108 | 108 | 108 | 108 |

**. Correlation is significant at the 0.01 level (1-tailed)

*. Correlation is significant at the 0.05 level (1-tailed)

**TABLE4:** Pearson's Product-moment Correlation Coefficients

Primarily, the proposed research model demonstrates high explanatory power, shown in Figure 4. The final path coefficient shows that the research model explains 48% of BI. The R-square of the Behavioral Intention (BI) from: Perceived Usefulness (PU) is .45; Perceived Ease Of Use (PEOU) is .34. The final result shows that all three determinants, PU, PEOU and PE significantly affect BI in mobile multimedia use.

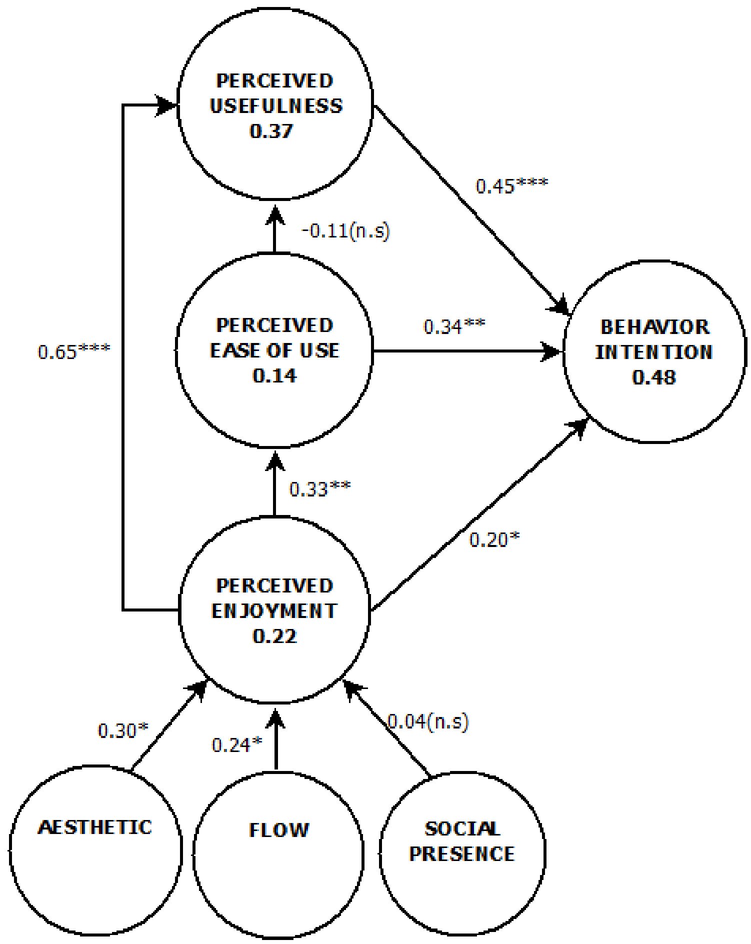



**FIGURE 4:** PLS results of the structural model

As for three given predictors towards PE, aesthetic and flow also show significant effects with 30% and 24% respectively. However, the PLS result does not show the significant effect of social presence as a predictor of PE and probable reasons will be discussed.

| Hypothesis | From | To | Beta | t-Value | p-Value | Sig | Status |
|---|---|---|---|---|---|---|---|
| H1 | PE | PU | 0.65 | 7.42 | 0.0000 | *** | Supported |
| H2 | PE | BI | 0.20 | 1.98 | 0.0251 | * | Supported |
| H3 | PE | PEOU | 0.33 | 3.58 | 0.0003 | ** | Supported |
| H4 | Aesthetic | PE | 0.30 | 2.51 | 0.0007 | * | Supported |
| H5 | Flow | PE | 0.24 | 2.14 | 0.0173 | * | Supported |
| H6 | Social Presence | PE | 0.04 | 0.47 | 0.3197 | - | n.s |

*p<0.5, **p<0.001, ***p<0.0001 (1-tailed)

**TABLE 5:** Result of hypotheses

The result of all hypotheses is listed in Table 5. The result showed that all hypotheses were supported except H6. Reviewing the results above, two sets of conclusions can largely be drawn. First, PE was positively associated with PU ($\beta = 0.65$; $p < 0.001$) and PEOU ($\beta = 0.33$; $p < 0.05$). Especially, PE had a strong association with PU. There was a relatively weak relationship between PE and BI and it can be assumed that PE had an effect through PU and PEOU toward BI, rather than direct effect on BI. One probable reason for this weak association of PE and BI is because of the diversity of mobile multimedia services. For instance, an activity to search information does not provide much perceived enjoyment, contrary to listening music or mobile blogging. Also, there seems to be moderating effect of PE. Hill and Troshani [105] studied Adoption of Personalization Mobile Services with young Australians and they also found that PE and PU were considered as important determinants of adoption in personalized ringtones. This is consistent with previous studies in the mobile information and entertainment services [106, 107]. Hill and Troshani [105] also suggested that mobile service adoption may vary depending on the type of services.

Second, aesthetics ($\beta = 0.30$; p-value $< 0.1$) and flow ($\beta = 0.24$; p-value $< 0.05$) show strong supports to explain PE. In the same context with why PE has a weak association with BI, social presence (SP) may have the same issues of types of mobile multimedia services. For instance, using SMS or emails requires higher SP, however taking self-photos with mobile phone is not related to social presence. All despondences of this study are based on the recent-used mobile multimedia services. Therefore there might be differences to explain PE, based on types of mobile multimedia services.

## 6. LIMITATIONS AND FUTURE RESEARCH

Limitations of this research provide the foundations for further studies to improve our understandings of the role of enjoyment and its determinants. Limitations may require cautious interpretations of this study. This study focuses on only a subset of possible determinants of mobile multimedia usage. However, mobile multimedia services are relatively lacking in consensus of its wide definitions and understandings. Also, there are wide variations of mobile multimedia services and it is hard to be covered by one criteria. As the result of this study shows, there might be more careful categorization for types of services.

This study employs only undergraduate student sample at a large Mid-Western university and the sample of this study cannot be generalized beyond its age group or its region. Furthermore, the female sample is three times larger than male group in this study and this also can lead different results when adopting other population. The self-reported recall measures of using mobile multimedia services also may be limitation of the present study. According to Pepper, Holmes, and Popovich (2004), direct observations can be more accurate measure than self-reports because people tend to underestimate their usage.

Findings of this study support most of hypotheses about the role of Perceived Enjoyment (PE) in mobile multimedia use. PE has an effect on Perceived Usefulness (PU), Perceived Ease Of Use (PEOU) and directly Behavior Intention (BI), even though the relationship with BI is relatively weak. Therefore, further research need to deeply examine the role of PE. PE may or may not be a dominant variable but moderating variable, and it may also depend on the context of use, types of services, motivation or situation.

The relationship between PE and aesthetics, flow and social presence should be further investigated. Even though, this study shows little relationship between them, there might be unobserved and overlooked indicators. Further research needs to examine alternative measures for aesthetics, flow and social presence with a specific context of use. Particularly for flow, as mentioned previously, this item includes multi-dimensions such as concentration, playfulness, time distortion, telepresence, and exploratory behavior [96]. Hence, it should be more clearly defined which sub-category is the optimal for measuring mobile multimedia services.

Additionally, mobile multimedia services are very diverse. The activity to take a photo using phone camera (i.e., self photo taking) is different from sending e-mails. There is little room for social presence when listening music or watching videos. Therefore, measuring social presence should be defined in activities with others. Hence, further research should consider the genres of mobile multimedia services. It also remains uncertain whether the relationships are unobserved. For instance, gender difference was observed. However, it cannot be further developed because the numbers of male respondents were relatively small. There were also possibilities of the relationship between mobile multimedia services and frequency of use. Therefore, further research is needed for these issues.

## 7. REFERENCES


[1] L. Manovich. *"Interaction as an aesthetic even"*. (2006) from Vodafone, Receiver magazine #17: Available at: http://www.vodafone.com/flash/receiver/17/articles/pdf/17_09.pdf, [Accessed 8 October 2008]

[2] J. Kim, J. Y. Moon. *"Designing towards emotional usability in customer interfaces — trustworthiness of cyber-banking system interfaces"*. Interacting with Computers, 10(1):1-29, 1998

[3] D. Cyr, M. Head, A. Ivanov. *"Design aesthetics leading to m-loyalty in mobile Commerce"*. Information and Management, 43(8):950-963, 2006

[4] P. Jordan. *"Designing pleasurable products"*. Taylor and Francis, (2000)

[5] S. Mahlke. *"Usability, Aesthetics, Emotions and the User Experience"*. Hot Topics, 6(2): 2007

[6] F. D. Davis. *"Perceived usefulness, perceived ease-of-use, and user acceptance of information technology"*. MIS Quarterly, 13(3):319–340, 1989

[7] J. Short, E. Williams and B. Christie. "*The social psychology of telecommunications"*, Wiley, (1976)

[8] M. Csikszentmihalyi. *"Flow: The psychology of optimal experience"*, Harper Perennial, (1991)

[9] ISO DIS 9242-11, (1998) Available at: http://www.iso.org,

[10] W. S. Green, P. W. Jordan. *"Human Factor in Product Design: Current Practice and Future Trends"*, Taylor & Francis, (1999)

[11] J. Nielsen. *"Usability Engineering"*, Academic Press (1993)

[12] J. Nielsen. *"Designing Web Usability: The Practice of Simplicity"*, New Riders Publishing, (2000)

[13] A. Dillon. "*Beyond usability: process, outcome and affect in human-computer interactions"*. Canadian Journal of Library and Information Science, 26(4):57-69, 2002

[14] P. Thomas, R. D. Macredie. *"Introduction to the new usability"*. ACM Transactions on Computer-Human Interaction (TOCHI), 9(2):69-73, 2002

[15] N. McNamara, J. Kirakowski. *"Functionality, usability, and user experience: three areas of concern"*. Interactions, 13(6):26-28, 2006

[16] M. G. Jones. *"Learning to Play, Playing to Learn: Lessons Learned from Computer Games"*. Paper presented at the Association for Educational Communications and Technology, Albuquerque, New Mexico, 1997

[17] M. Keil, P. M. Beranek and B. R. Konsynski. *"Usefulness and ease of use: field study evidence regarding task considerations"*. Decision Support Systems, 13(1):75-91, 1995

[18] Y. Malhotra, F. Galletta. *"Extending the technology acceptance model for social influence: theoretical bases and empirical validation"*. In Proceedings of the 32th International Conference on System Sciences Los Alamitos, CA, 1999

[19] J. W. Moon, Y. G. Kim. *"Extending the TAM for a world-wide-web context"*. Information & Management, 38(4):217–230, 2001

[20] D. A. Adams, R. R. Nelson and P. A. Todd. *"Perceived usefulness, ease of use, and usage of information technology: a replication"*. MIS Quarterly, 16(2):227-247, 1992

[21] W. C. Chin, P. A. Todd. *"On the use, usefulness and ease of use of structural equation modeling in MIS research: a note of caution"*. MIS Quarterly, 19(2):237-246, 1995

[22] W. J. Doll, A. Hendrickson and X. Deng. *"Using Davis's perceived usefulness and ease-of-use instruments for decision making: a confirmatory and multi-group invariance analysis*". Decision Sciences, 29(4):839-869, 1998

[23] K. Mathieson. *"Predicting use intentions: comparing the technology acceptance model with the theory of planned behavior"*. Information Systems Research, 2(3):173-191, 1991

[24] A. H. Segars, V. Grover. *"Re-examining perceived ease of use and usefulness: a confirmatory factor analysis*". MIS Quarterly, 17(4):517-525, 1993

[25] P. Vorderer, C. Klimmt and U. Ritterfeld. *"Enjoyment: At the Heart of Media"*. Communication Theory, 14(4):88-408, 2004

[26] N. McNamara, J. Kirakowski. *"Defining usability: quality of use or quality of experience?"*. In Proceedings of IEEE International Professional Communication Conference, Limerick, Ireland, 2005

[27] J. Ollila. *"Structuring for Growth".* Paper presented at Nokia Capital Market Days, New York, USA, 2003

[28] T. Ojala, J. Korhonen J, M. Aittola, M. Ollila, T. Koivumäki, J. Tähtinen and H. Karjaluoto. *"SmartRotuaari - Context-aware mobile multimedia services"*. In Proceedings of the 2nd International Conference on Mobile and Ubiquitous Multimedia, Norrköping, Sweden, 2003

[29] G. D. Abowd, C. G. Atkeson, J. Hong, S. Long, R. Kooper and M. Pinkerton. *"Cyberguide: A mobile context-aware tour guide"*. ACM Wireless Networks, 3:421–433, 1997

[30] H. Verkasalo. *"Empirical observations on the emergence of mobile multimedia services and applications in the U.S. and Europe"*. In Proceedings of the 5th international conference on Mobile and ubiquitous multimedia, Stanford, CA, USA, 2006

[31] F. D. Davis, R. P. Bagozzi and P. R. Warshaw. *"Extrinsic and intrinsic motivation to use computers in the workplace"*. Journal of Applied Social Psychology, 22(14):1111–1132, 1992

[32] R. M. Ryan, E. L. Deci. *"Intrinsic and extrinsic motivations: Classic definitions and new directions"*. Contemporary Educational Psychology, 25(1):54–67, 2000

[33] M. Y. Yi, Y. Hwang. *"Predicting the use of web-based information systems: self-efficacy, enjoyment, learning goal orientation, and the technology adoption model"*. International Journal of Human Computer Studies, 59(4):431–449, 2003

[34] H. Sun, P. Zhang. *"An Exploration of Affect Factors and Their Role in User"*. Journal of the American society for information science and technology, 59(8):1252-1263, 2008

[35] R. Agarwal, E. Karahanna. *"Time flies when you're having fun: cognitive absorption and beliefs about information technology usage"*. MIS Quarterly, 24(4):665-694, 2000

[36] V. Venkatesh. "*Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model"*. Information Systems Research, 11(4):342-365, 2000

[37] V. Venkatesh, C. Speier and M .G. Morris. *"User acceptance enablers in individual decision making about technology: toward an integrated model".* Decision Sciences, 33(2):297–316, 2002

[38] J. H. Huang, Y. R. M. Lin and S. T. Chuang. *"Elucidating user behavior of mobile learning"*. The Electronic Library, 25(5):586-599, 2007

[39] V. L. Blakney, W. Sekely.*"Retail Attributes: Influence on Shopping Mode Choice Behavior"*. Journal of Managerial Issues, 6(1):101-118, 1994

[40] A. M. Forman, V. Sriram. *"The Depersonalization of Retailing: Its Impact on the 'Lonely' Consumer"*. Journal of Retailing, 67(2):226-243, 1991

[41] S. L. Jarvenpaa, P. A. Todd. *"Is There a Future for Retailing on the Internet?".* In Electronic Marketing and the Consume", R. A. Peterson (ed.), Sage., pp. 139-154 (1997)

[42] A. S. Dick, K. Basu. "*Customer Loyalty: Toward an Integrated Conceptual Framework".* Journal of the Academy of Marketing Science, 22(2):99-113, 1994

[43] M. Koufaris. "*Applying the Technology Acceptance Model and Flow Theory to Online Consumer Behavior"*. Information Systems Research, 13(2): 205–223, 2003

[44] D. Li, P. Y. K. Chau and H. Lou. *"Understanding individual adoption of instant messaging: An empirical investigation"*. Journal of the Association for Information Systems, 6(4):102-129, 2005

[45] J. Y. L. Thong, S. J. Hong and K. Y. Tam. "*The effects of post-adoption beliefs on the expectation-confirmation model for information technology continuance"*. International Journal of Human-Computer Studies, 64(9):799-810, 2006

[46] Z. Huang, J. J. Cappel.*"Assessment of a Web-Based Learning Game in an Information Systems Course"*. Journal of Computer Information Systems, 45(4):42-50, 2005

[47] K. H. Kim, J. Y. Park, D. Y. Kim, H. I. Moon and H. C. Chun. *"E-Lifestyle and Motives to Use Online Games"*. Irish Marketing Review, 15(2):71-77, 2002

[48] J. Wu, D. Liu. *"The effects of trust and enjoyment on intention to play online games"*. Journal of electronic commerce research, 8(2):128-140, 2007

[49] V. Venkatesh, V. Ramesh. *"Usability of web and wireless sites: extending the applicability of the Microsoft usability guidelines instrument"*. Information Systems Technical Reports and Working Paper, 2002

[50] D. E. Rosen, E. Purinton. *"Website design: viewing the web as a cognitive landscape"*. Journal of Business Research, 57(7):787–794, 2004

[51] Z. Jiang, I. Benbasat. *"The effects of interactivity and vividness of functional control in changing web consumers' attitudes"*. In Proceedings for the 24th International Conference on Information Systems, Seattle, USA, 2003

[52] D. Cyr, C. Bonanni, J. Bowes, J and J. Ilsever. *"Beyond trust: website design preferences across cultures"*. Journal of Global Information Management, 13(4):24–52, 2005

[53] D. Gefen, D. W. Straub D. W., and M .C. Boudreau. *"Structural equation modeling and regression: guidelines for research practice"*. Communications of the Association for Information Systems, 4(7):2–77, 2000

[54] V. Postrel. "*The substance of style"*, Harper Collins., (2002)

[55] S. Bly, L. Cook, T. Bickmore, E. Churchill and J. W. Sullivan. *"The rise of personal web pages at work"*. In Proceedings of the conference on human factors in computing systems, Los Angeles, CA, USA, 1998

[56] M. Hassenzahl. *"The thing and I: understanding the relationship between user and product"*. In M. A. Blythe, A. F. Monk, K. Overbeeke and P. C. Wright (eds.). Funology: From usability to enjoyment. Kluwer Academic Publisher., (2003)

[57] N. Tractinsky. *"Does Aesthetics Matter in Human Computer Interaction?"* Mensch & Computer, S:29-42, 2005

[58] H. Heijden. *"Factors influencing the usage of websites: the case of a generic portal in The Netherlands"*. Information & Management, 40(6):541–549, 2003

[59] M. Csikszentmihalyi. *"The future of flow"*. In M. Csikszentmihalyi, I. S. Csikszentmihalyi (Eds.), Optimal experience: Psychological studies of flow in consciousness, Cambridge University Press., pp. 364–383 (1988a)

[60] M. Csikszentmihalyi. *"The flow experience and its significance for human psychology"*. In M. Csikszentmihalyi, I. S. Csikszentmihalyi (Eds.), Optimal experience: Psychological studies of flow in consciousness, New York: Cambridge University Press., pp. 15–35 (1988b)

[61] J. L. Sherry. *"Flow and Media enjoyment"*. Communication Theory, 14(4):328-347, 2004

[62] M. Csikszentmihalyi. "*The evolving self: A psychology for the third millennium"*, Harper Perennial (1993)

[63] C. L. Hsu, H. P. Lu. *"Why do people play on-line games? An extended TAM with social influences and flow experience"*. Information & Management , 41(7):853-868, 2004

[64] M. G. Jones. *"Creating engagement in computer-based learning environments"*, ITFORUM (1998)

[65] K. Chen, M. J. Tarn, B. T. Han. *"Internet dependency: Its impact on online behavioral patterns in E-commerce"*. Human Systems Management, 23(1):49-58, 2004

[66] L. K. Trevino, J. Webster. *"Flow in computer mediated communication: Electronic mail and voice mail evaluation and impacts"*. Communication Research, 19(5):539–573, 1992

[67] J. Webster, L. K. Trevino and L. Ryan. *"The dimensionality and correlates of flow in human-computer interaction"*. Computers in Human Behavior, 9(4):411–426, 1993

[68] J. A. Ghani, S. P. Dashpande. *"Task characteristics and the experience of optimal flow in human-computer interaction"*. The journal of psychology, 128(4): 381–391, 1994

[69] D. L. Hoffman, T. P. Novak. *"Marketing in hypermedia computer-mediated environments: Conceptual foundations"*. Journal of Marketing, 60(3):50–117, 1996

[70] T. P. Novak, D. L. Hoffman and Y. F. Yung. *"Modeling the structure of the flow experience among Web users"*. Abstract for the INFORMS Marketing Science and the Internet Mini-Conference, 1998

[71] A. Mehrabia, J. A. Russel. *"An Approach to Environmental Psychology"*, MIT Press., (1974)

[72] Herz, J. C. *"Joystick Nation: How video games ate our quarters, won our hearts, and rewired our minds"*, Little Brown and Company (1997)

[73] Y. Lu, T. Zhou, B. Wang. *"Exploring Chinese users' acceptance of instant messaging using the theory of planned behavior, the technology acceptance model, and the flow theory*". Computers in Human Behavior, 24(1):29-39, 2009

[74] J. Fulk, J. Schmitz, G. J. Power. *"A Social Information Processing Model of Media Use in Organizations"*. Communication Research, 14(5):520-552, 1987

[75] F. Biocca, C. Harms, J. K. Burgoon. *"Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria"*. PRESENCE, 12(5):456-480, 2003

[76] K. Hassanein, M. Head. *"The Impact of Infusing Social Presence in the Web Interface: An Investigation Across Product Types"*. International Journal of Electronic Commerce, 10(2):31-55, 2006

[77] J. Connell, G. Mendelsohn, R. Robins and J. Canny. *"Effects of communication medium on interpersonal perceptions"*. In Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work, Boulder, CO, USA, 2001

[78] P. L. P. Rau, Q. Gao and L. M. Wu. *"Using mobile communication technology in high school education: Motivation, pressure, and learning performance*". Computers & Education, 50(1):1-22, 2008

[79] M. Heilbronn, W. Libby. *"Comparative effects of technological and social immediacy upon performance and perceptions during a two-person game"*. Paper presented at the annual convention of the American Psychological Association, Montreal, Canada, 1973

[80] K. Cumiskey. *"Mobile fantasies on film: Fostering "co-presence" through the integrated use of real and imagined mobile technology"*. Pre-Conference, International Communication Association, 2006

[81] N. Eagle, A. Pentland. *"Social serendipity: Mobilizing social software".* IEEE Pervasive Computing, 4(2):28-34, 2005

[82] L. Humphreys. *"Mobile Devices and Social Networking"*. Pre-Conference, International Communication Association, 2006

[83] I. E. Smith. *"Social mobile applications"*. IEEE Computer Society, 38(4):84-85, 2005

[84] H. Tomita. *"Geitai and the Intimate Stranger"*. In M. Ito, D. Okabe and M. Matsuda (eds.). Personal, Portable, Pedestrian. The MIT Press., (2005)

[85] S. H. Kang, J. H. Watt and S. K. Ala. *"Social Copresence in Anonymous Social Interactions"*. In Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing, Florence, Italy, 2008

[86] R. Ling. *"The mobile connection: The cell phone's impact on society"*. Morgan Kaufmann Publishers., (2004)

[87] K. Nyiri. *"Mobile Democracy: Essays on society, self and politics"*, Passagen Verlag., (2003)

[88] S. L. Jarvenpaa, K. R. Lang, Y. Takeda, V. K. Tuunainen. *"Mobile Commerce At Crossroads"*. Communications of the ACM, 46(12):41-44, 2003

[89] C. Heeter. *"Communication research on consumer VR"*. In F. Biocca, M. R. Levy (eds.), Communication in the age of virtual reality, Lawrence Erlbaum Associates., pp191-218 (1995)

[90] M. Lombard, T. Ditton. *"At the heart of it all: The concept of presence"*. Journal of Computer Mediated Communication, 3(2): 1997

[91] K. Hassanein, M. Head. *"Manipulating perceived social presence through the web interface and its impact on attitude towards online shopping"*. International Journal of Human-Computer Studies, 65(8):689-708, 2007

[92] MarketWire. *"Mobile Tuesday to Deliver Coupons, Savings via Mobile Phones Every Tuesday for 52 Week"*. Available at : http://www.istockanalyst.com/article/viewiStockNews/articleid/2866954 [Accessed 8 October2008]

[93] FuorDigital. *"Research Brief: Mobile Phone Marketing"*. Available at: www.fuor.net/dnn/Portals/0/Research_Brief_Mobile_Phone_Marketing.pdf [Accessed 8 October 2008]

[94] M:Metrics. "*Teen, College Students Are Most Active Cell Phone Users"*. (2005) from (http://www.clickz.com/3530886), [Accessed 8 October 2008]

[95] H. Geser. *"Towards a Sociological Theory of the Mobile Phone"*. from Swiss Online Publications in Social Sciences: Available at: http://www.itu.dk/people/ldn/Lars/Geser_lektion6.pdf [Accessed 8 October 2008]

[96] T. J. Chou, C. C. Ting. *"The Role of Flow Experience in Cyber-Game Addiction"*. Cyber Psychology & Behavior, 6(6):663-675, 2003

[97] H. Nysveen, P. E. Pedersen and H. Thorbjornsent. *"Explaining intention to use mobile chat service: moderating effects of gender"*. The journal of consumer marketing, 22(5):247–256, 2005

[98] J. Yu, I. Ha, M. Choi and J. Rho. *"Extending the TAM for a t-commerce"*. Information & Management, 42(7):965-976, 2005

[99] M. Soderlund, N. Ohman. "*Behavioral Intentions in Satisfaction Research Revisited"*. Journal of Consumer Satisfaction and Dissatisfaction and Complaining Behavior, 16:53-66, 2003

[100] T. Garling, S. Fujii, A. Garling, A and C. Jakobsson. *"Moderating effects of social value orientation on determinants of proenvironmental behavior intention"*. Journal of Environmental Psychology, 23(1):1-9, 2003

[101] D. Gefen, D. W. Straub. *"Consumer trust in B2C e-Commerce and the importance of social presence: experiments in e-Products and e-Services"*. Omega, 32(6):407-424, 2004

[102] V. Venkatesh. *"Creation of favorable user perceptions: Exploring the role of intrinsic motivation"*. MIS Quarterly, 23(2):239-260, 1999

[103] J. Éthier, P. Hadaya, J. Talbot and J. Cadieux, *"B2C Web Site Quality and Emotion During Online Shopping Episodes: An Empirical Study"*. Information and Management, 43:627-639, 2006

[104] W. W. Chin. *"The Partial Least Squares approach for structural equation modeling",* Lawrence Erlbaum Associates., pp.295-336 (1998)

[105] S. R. Hill, I, Troshani. "*Adoption of Personalisation Mobile Services: Evidence from Young Australians"*. 22nd Bled eConference eEnablement: Facilitating an Open, Effective and Representative eSociety, Bled, Slovenia, 2009

[106] D. J. Xu. *"The influence of personalization in affecting consumer attitudes toward mobile advertising in China"*. Journal of Computer Information Systems, Winter 2006-2007:9-19, 2007

[107] F. B. Tan, J. P. C. Chou. "*The relationship between mobile service quality, perceived technology compatibility, and users' perceived playfulness in the context of mobile information and entertainment services*". International Journal of Human-Computer Interaction, 24(7):649-671, 2008

# Real Time Blinking Detection Based on Gabor Filter

**Kohei Arai** arai@is.saga-u.ac.jp
*Information Science*
*Saga University*
*Saga, 840-0027, Japan*

**Ronny Mardiyanto** ronny_mardiyanto@yahoo.com
*Electrical Engineering*
*Institut Teknologi Sepuluh Nopember*
*Surabaya, Indonesia*

**Abstract**

New method of blinking detection is proposed. The utmost important of blinking detection method is robust against different users, noise, and also change of eye shape. In this paper, we propose blinking detection method by measuring the distance between two arcs of eye (upper part and lower part). We detect eye arcs by apply Gabor filter onto eye image. As we know that Gabor filter has advantage on image processing application since it able to extract spatial localized spectral features such as line, arch, and other shapes. After two of eye arcs are detected, we measure the distance between arcs of eye by using connected labeling method. The open eye is marked by the distance between two arcs is more than threshold and otherwise, the closed eye is marked by the distance less than threshold. The experiment result shows that our proposed method robust enough against different users, noise, and eye shape changes with perfectly accuracy.

**Keywords:** Blinking, Gabor Filter, Eye Arc.

## 1. INTRODUCTION

Blinking as one of eye movement activity beside glance has been used to keep the moisture of eye. Blinking helps eye to spread the tears and also remove unwanted things from eye. This activity is controlled by brain automatically (voluntary). When eye become dry or there is unwanted things on eye, brain will send command to eye muscle in order to blink the eye. The blinking activity is also affected by fatigue, disease, drowsiness, and etc. That is why blinking also has been used as one common parameter for measuring fatigue and drowsiness. The drowsy driver can be measured based on the blinking rhythm. Vehicle safety system based on drowsiness parameter is proposed [3]. This system determines the level of drowsiness by using blinking. When it is used as drowsiness parameter, the accuracy become important because it should be able to detect not only involuntary blinking, but also voluntary blinking which is has short duration period. Beside this biological function, blinking also can be used to show the user attention. Using blinking, dumb people can communicate with others. Aid communication based on blinking has been proposed [4]. The system enables communication using "blink patterns", sequences of long and short blinks, which are interpreted as semiotic messages. Also, such Morse code can be used as one of language tool for communication among dumb peoples. Moreover, Human Computer Interaction Application has been using blinking to help the selection process [5]. User can type on computer by utilize blinking only. Such false/true choice can be selected by user using blinking only. The method of blinking detection can be broadly classified into following:

1. Biological Approach, by using EOG [1] (Attach the surface electrode onto surrounding eye. Blinking is detected by measuring eye muscle potential in vertical and horizontal direction) or EEG [2] (Attach the surface electrode into skull surface in order to measure brain activity),
2. Image Analysis: Capture eye image by using camera. Several image-processing steps are needed to observe the blinking.

The method (1) is relative expensive and not convenience compared than method (2) since method (1) burden the user as several electrodes have to be attached onto user's skin. In this paper we focus on blinking detection based on image analysis approach. Among the blinking application, the utmost important of blinking detection method is accurate against eye shape changes, varies of blinking speed, varies of users, and also noise. Ref [6] proposed blinking detection based on Hough Transform. Hough transform has been used to detect iris of opened eye. Refs [7][9][10][14] proposed blinking detection based on the motion. Such normal flow and optical flow have been used to detect the blinking. Ref [7] detected blinking based on iris of opened eye. By detect eyelids as reference points first, opened and closed eye are estimated from its points. Normal flow method has been used to improve the accuracy of iris detection. Ref [14] used combination between normal flow and deterministic finite state machine to improve the detection accuracy. Ref [8] detected blinking by variance map of sequential frames and threshold. Ref [11] detected blinking by using opened eye template. Ref [12] detected blinking by using Adaboost Classifier. Ref [13] detected blinking by using Stereo Vision.

Ref [6] proposed method for identifying the opened and closed state of eye in real time video. First, they separated the eye region then applied Hough transform to detect circle on eye which represent iris. This method has 94% accuracy on eye state detection. Ref [7] located eye corners, eyelids, and irises of each frame and analyzed its motion to detect the change of gaze direction and blinking. By using simple model of head and eye, they determine the head-independent motions of the irises and eyelids. They detect blinking by tracking upper eyelid and measuring distance between its apex and center of iris. The normal flow has been used to track all features points. Ref [9] proposed blinking detection for eye fatigue measurement. They recorded video of user by using head mounted camera. First, they separate face image then continue with detect the eyes. Eyes locations are also detected by optical flow. Blinking is detected by using normal flow and adaptive threshold. Ref [10] implemented GPU on real-time eye blink detector. After eyes locations have been detected, eye is separated as region of interest. By using optical flow, closed eye is marked when the dominant motion is downward. Otherwise, opened eye is marked when the dominant motion is upward. Ref [14] used combination between boosted classifier and Lucas-kanade for tracking the face and eyes positions. Basically, blinking is detected by using normal flow. In order to improve accuracy, discrete finite state machine is used. Ref [8] used spatio-temporal filtering to locate the head position in an image sequence. The located head position is tracked and use variance map to identify any possible eye blink pair candidates. After eye-blink pair is detected, the contours around are adjusted and record four feature points for each eye. These feature points are tracked by using modified version of Lucas-Kanade method. Ref [11] detected user's eye blink and analyzed the pattern and duration to provide input to computer input as a mouse click. Eye is tracked in real time using correlation with online template. Opened eye template is used to find and track the eye position. Blinking is estimated by comparing the similarity between opened eye template and current image. When user closes the eye during blinking process, the similarity will decrease. Otherwise, the similarity will maximum when user fully open the eye.

Among previous methods, the disturbance such as noise, varies of eye shapes, and varies of blinking speed still challenge to be solved. Hough transform-based has problem against noise and varies of eye shapes. Normal flow and Optical flow have advantage robust against varies of eye shape, but its have weakness against varies of blinking speed. When it is operated in normal speed, optical/normal flow looks success detects the blinking. Unfortunately, when it is operated very slow or very fast condition, accuracy become decrease (If it is too slow, no motion will be detected. Likewise, if it is too fast, motion will hard to be captured). Template matching based has problem against varies of eye shapes (varies of users).

The objective of this research is to overcome different users, noise, and accuracy problem of previous blinking detection methods. In this paper we propose blinking detection method by utilizing Gabor filter. Gabor filter extract the eye arcs and then measure the distance between these arcs to determine the blinking. Web camera mounted on top of display is used to acquire the image. First, eye location is detected by using adaboost classifier (proposed by Viola-Jones). The advantage of this method is fast and robust against different users. After Adaboost classifier success detects the eye location, eye is cropped as region of interest (ROI). To avoid user's movement problem, the stabilizer of ROI image is used. The facial feature points such as eye corners and eye lids are used to stabilize the ROI by applying affine transformation. Affine transformation and the features points guarantee that the preprocessing image has same size and rotation angle between each frame. After the stable eye image has been obtained, we apply Gabor filter into account. Gabor filter extracts the pattern of eye based on the orientation angle. Gabor filter can distinguish the orientation angle of each line on eye. Because of this, Gabor filter can separate between top arc and bottom arc of eye based on the rotation angle (because both of arcs are close with horizontal lines, so its can be separated by using orientation angle is 0º). By using connected labeling method, the positions of both arcs are easily detected. Blinking is detected by measuring the distance between these two arcs. If this distance is more than threshold then opened eye is detected. Otherwise, if the distance is less than threshold then closed eye is detected. We made experiments to measure the accuracy of our proposed method by involving several users. Users blink the eyes in different gaze directions to prove that our proposed method robust against varies of eye shapes (different gaze directions make eye shape looks different on each). Moreover, we compare the accuracy of our proposed method with well-known methods such as opened eye template and normal flow methods.

This paper is organized as follows: section 2 describes our proposed system involving hardware configuration, eye detection and tracking, ROI stabilizer, and blinking detection by using Gabor filter. Section 3 describes our experimental results involving blinking detection performances against noise, varies of eye shape, and accuracy. Section 4 describes the conclusion.

## 2. Proposed Method

In this paper we propose new blinking detections method by using Gabor filter. Gabor filter extracts the pattern of eye to top and bottom arcs. While two arcs have been detected, the distance between both is measured and used to determine the blinking.

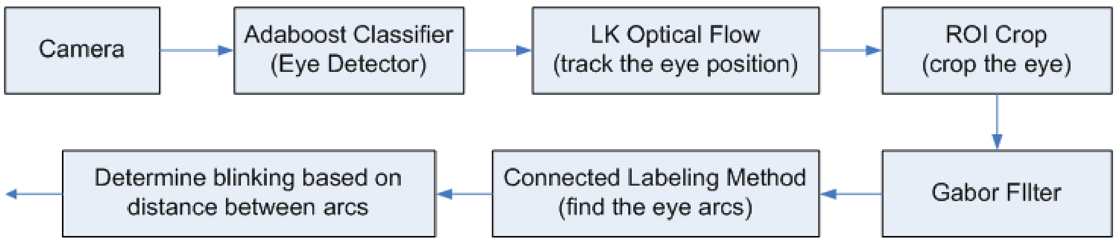



**FIGURE 1:** Block diagram of proposed blinking detection method

In our blinking detection method, we use IR camera as main source for acquiring eye image. The light sensor on camera is covered by the tape to make IR LEDs always ON. This way is worthwhile to maintain the illumination on eye image and eliminate the illumination changes and yield the stable eye image. After the stable eye image is obtained, Adaboost Classifier is used to detect the eye position for the first time only. After the eye location is founded, the next eye image is obtained from the use of Lucas-Kanade optical flow by tracking the new position of eye. This obtains ROI of eye image and only this image is used to detect the blinking in next process. We apply Gabor filter onto this ROI to extract two of eye arcs. The locations of all arcs are detected by using Connected Labeling method. After all locations of arcs have been detected, the closed eye is marked when the distance between top and bottom arcs is below than threshold. Otherwise, the opened eye is marked when the distance is above than threshold. The block diagram of entire process is shown in Figure 1.

### 2.1 Hardware Configuration

We setup our hardware configuration as depicted in Figure 2. The NetCowBow DC-NCR 131 IR camera, 1.3 Mega pixels with eight IR LEDs, is used as input. The IR LED will compensate every the illumination changes. By using this type of camera, the preprocessing image will be robust against the illumination changes (we will get the stable and robust image against illumination changes). The Camera is placed on top of display as depicted in Figure 2. The distance between user and display is 30cm. We use Optiplex 755 dell computer with Core 2 Quad 2.66 GHz CPU and 2G RAM. We develop our software using C++ Visual Studio 2005 and open source of OpenCv Image processing Library.

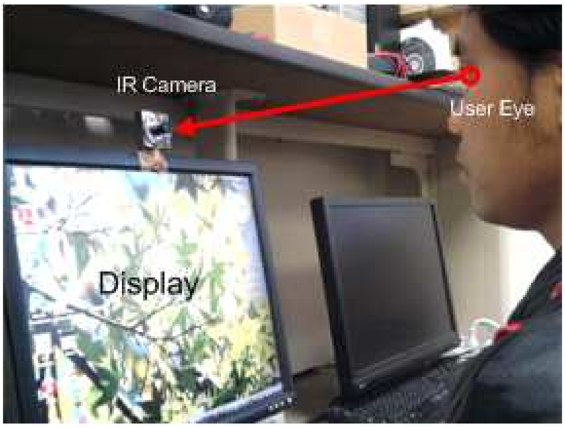



**FIGURE 2:** Hardware Configuration.

### 2.2 Eye Detection and Tracking

Before detect the blinking, it is very important to prepare the stable eye image. Because the positions of users always change, the position of eye also may change due to this. The user's movement may affect the stability of image because the ROI of eye always change. Unstable image will decrease accuracy of blinking detection. Also, it must guarantee that the eye image must have same size, heads rotation, and symmetry between right and left. Head rotation means that the rotations value of eye image must same even if the user rotate their head. In order to obtain the good and stable eye image, we detect eye location first. Eye location is detected by utilize Adaboost Classifier (One of OpenCv function). In order to detect eye location by using Adaboost Classifier, the following code is used,

```
cvHaarDetectObjects(small_img,cascade,storage,1.1,2,0|CV_HAAR_FIND_BIGGEST_OBJECT,cvSize(30,30));
```

By using the above code, the eye location will fast and easily detected. After once eye location is detected, next is tracking this eyes location by using Lucas-Kanade optical flow. Because LK optical flow needs good feature points to tract the eye location, we use eyebrows as tracked points. The new eye location is estimated based on these tracked points.

### 2.3 Gabor-based blinking detection

Since Gabor [15] introduced Gaussian-modulated complex exponentials and Daugman [16],[17] generalized it to 2D form, 2D Gabor has been well-know method for analyzing the pattern. Also the used of Gabor wavelet to represent the image has been introduced by Tai Sing Lee [18]. In this paper, we use 2D Gabor filter to extract the eye pattern based on their angle. Before we extract the eye pattern, we broadly classify eye image into (1) line (such as eye lids, eye arcs) and (2) circle (such as iris, pupil). We will use Gabor filter to separate the eye arcs from other eye`s components and then estimate its positions.

Typically, we can extract the pattern by using edge detector such as sobel, canny, and etc. Unfortunately, the edge detector will not give information about their angle. Also, it cannot distinguish between line and bar. Because of this, we use Gabor filter to separate arcs from others eye's component and estimate its locations. The 2D Gabor can be written as equation (1) [19].

$$G(x,y) = \exp\left(\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\cos(2\pi\frac{x`}{\lambda}+\varphi)\right) \quad (1)$$

Where,

$x' = x\cos\theta + y\sin\theta$

$y' = -x\sin\theta + y\cos\theta$

In equation (1) we can see that it compose from Gaussian function (with σ is variance modulated with sinusoidal function (with λ is wavelength). $\theta$ is orientation, φ is phase offset, γ is spatial aspect ration.

In our system, we use Gabor filter by creating Gabor kernel. We create Gabor kernel by set the variance, wavelength, orientation, phase offset, and spatial aspect ratio. The example of Gabor kernel is depicted in Figure 3. This kernel contains 2D array values obtaining from the calculation using Gabor equation.

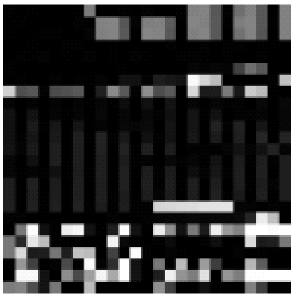


**FIGURE 3:** Example of 25 by 25 of Gabor kernel.

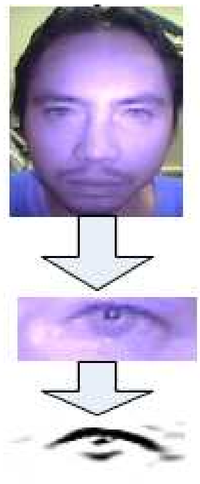


**FIGURE 4:** Process of converting from image source to the eye arcs image

In this paper we use Gabor filter to separate the arcs from all unnecessary eye features such as skin color, eyelid, eyebrow, iris, pupil, sclera, and etc. it will eliminate unused eye's component but keep the outer shape of eye. So the result image is line/arcs which represent the outer shape of eye (remove iris, pupil, and etc). The process from original source image to arcs image by using Gabor filter is shown in Figure 4.

After eye arcs image is created, next is estimation of the arcs locations. In order to measure the arcs length, we utilize connected component labeling method. This method estimate the arcs location based on their connectivity. Arcs are grouped into components and counted the amount of pixels. We estimate the top and bottom arcs based on their length, highest and lowest position, and relation between the arcs itself. By using this way, the noise will be automatically eliminated. After the arcs locations are found, the final value of blinking is estimated by using following rules:
**If** distance < threshold **then** closed eye detected
**If** distance >= threshold **then** opened eye detected
**If** opened eye detected -> closed eye detected -> opened eye detected **then** blinking detected

## 3. EXPERIMENT RESULTS

To measure the performance of our proposed method, the experiments have been done by involving three different users. Each user has to blink their eyes in three different gaze direction conditions (forward, left, and right) as depicted in Figure 5. This kind of experiment will prove that our proposed method does work when the user's gaze direction changes. As we can see in Figure 6, Figure 7, and Figure 8, the images change when the users look at different directions. When the users change their gaze directions, the eye shapes change because the eye ball rotate and iris (black eye) also change. This situation makes eye shape change and may affect the accuracy of blinking detection.

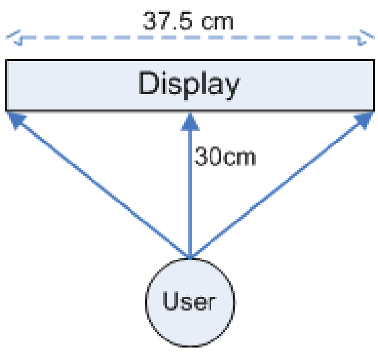
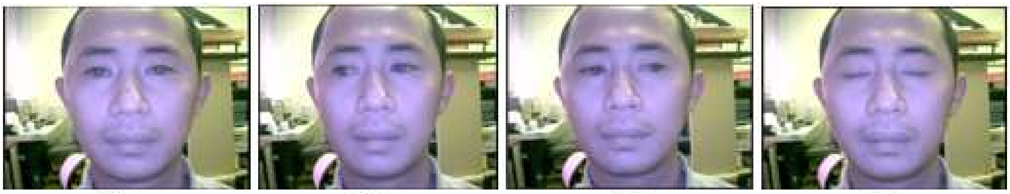



**FIGURE 5:** User looking at three different gaze directions

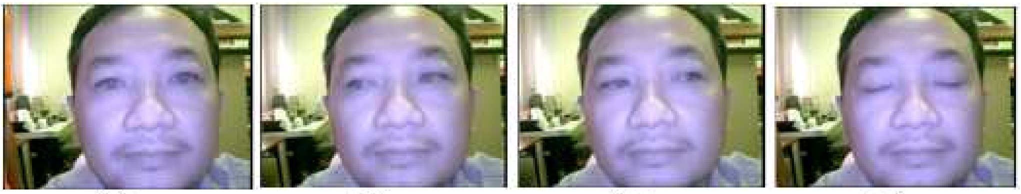
The taken images from three users are shown in Figure 6, Figure 7, and Figure 8. The number of images of user 1, user 2, and user 3 are 895, 567, and 609 images. The user's images look blue because it was obtained from IR camera.


**FIGURE 6:** Taken images from user 1, (a) looking forward, (b) looking at right, (c) looking at left, and (d) blink



(a) (b) (c) (d)

**FIGURE 7:** Taken images from user 2, (a) looking forward, (b) looking at right, (c) looking at left, and (d) blink



(a) (b) (c) (d)

**FIGURE 8:** Taken images from user 3, (a) looking forward, (b) looking at right, (c) looking at left, and (d) blink

### 3.1 Accuracy

The first experiment investigates the accuracy and variance of proposed method against varies users (varies eye shape) and different gaze directions. First, we counted all blinking from the video sources manually. The total blinking is used as the reference correct number of blinking. After we counted the total of all blinking, we counted the blinking by using our proposed method, template matching method, and motion template method.

In template matching method, we use opened eye image as template which is taken from one of users. The similarity between template image and actual eye image are compared in order to determine the blinking. One template is used to detect all users. In motion template method, we use motion template functions in OpenCv. These functions will give use information about the magnitude and angle of motion which is happen on image. We detect the blinking if the magnitude is more than threshold and the angle is around 90° and 270° (Upward and downward directions which represent the blinking activity (closed eye is estimated by detect the downward direction (angle is 270°) and opened eye estimated by detect the upward direction (angle is 0°))).

The accuracy and variance against different users is shown in table.1. Table.1 shows that our proposed method superior than others because our method didn't use direct image for blinking detection. The use of Gabor filter to extract eye arcs and continue with the use of connected labeling method to measure the distance between the arcs is effective for blinking detection because even the shape of eye changes (due to noise or different users), Gabor filter still success extract the arcs and estimate the blinking. Table 1 shows that our proposed method accurate against different users and varies of eye shape (with 100% accuracy and the variance against different users is zero). As we can see that the other methods only have accuracy 71.18 % and 83.75% and also the variances are 720.15 and 74.13. The accuracy of template matching method becomes decrease against different users because the template image only appropriate for one user. If this template is used for other users, the similarity becomes decrease. The motion template method looks more powerful than template matching method because they detect from the users motion. Unfortunately, if users blink their eye with slow speed or very fast speed the accuracy becomes decrease. This data proved that our proposed method robust against different users and varies of eye shape.

| | Method | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Our Method | | Template Matching | | Motion Template | |
| User | Total Blinking | Detected | Accuracy (%) | Detected | Accuracy (%) | Detected | Accuracy (%) |
| User 1 | 32 | 32 | 100 | 31 | 96.88 | 26 | 81.25 |
| User 2 | 30 | 30 | 100 | 22 | 73.33 | 23 | 76.67 |
| User 3 | 30 | 30 | 100 | 13 | 43.33 | 28 | 93.33 |
| **Average** | | | **100** | | **71.18** | | **83.75** |
| **Variance** | | | **0** | | **720.15** | | **74.13** |

**TABLE 1:** Accuracy and variance against different users

### 3.2 Noise Influence

In order to measure the performance of our proposed method against noise, we add Gaussian noise onto source image. The Gaussian noise with mean value is zero and adjustable standard deviation is added onto original image (8 bit/ gray image). We used OpenCv function to create the noise as follow,

```
CvRNG rng = cvRNG(-1);
IplImage* noise=cvCreateImage( cvSize(gray->width,gray->height),IPL_DEPTH_32F, 1 );
cvRandArr( &rng, noise, CV_RAND_NORMAL, cvScalarAll(mean), cvScalarAll(std_deviation) );
cvAcc( gray, noise );
cvConvert( noise, gray );
cvReleaseImage( &noise );
```

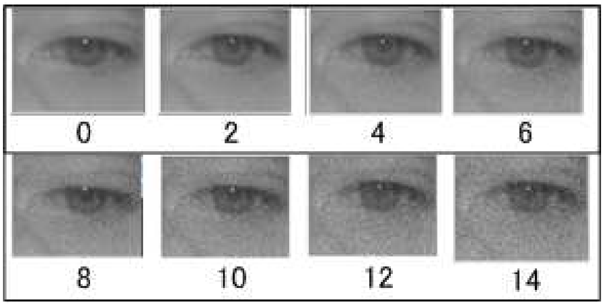



**FIGURE 9:** Added noise images of user 1 (standard deviation is 0, 2, 4, 6, 8, 10, 12, and 14)

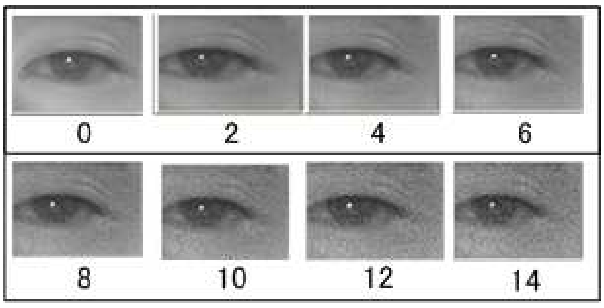



**FIGURE 10:** Added noise images of user 2 (standard deviation is 0, 2, 4, 6, 8, 10, 12, and 14)

In this experiment, we just add the noise (created by using Gaussian distribution with adjustable standard deviation) onto original image. The added noise image is shown in figure 9, 10, and 11. These figures show how noise can make eye images become not clear and decrease the accuracy.


**FIGURE 11:** Added noise images of user 3 (standard deviation is 0, 2, 4, 6, 8, 10, 12, and 14)

We used the added noise image of all users and apply our proposed method onto these. We recorded accuracy while noise is added (with varies values) of each user as shown in figure 12.

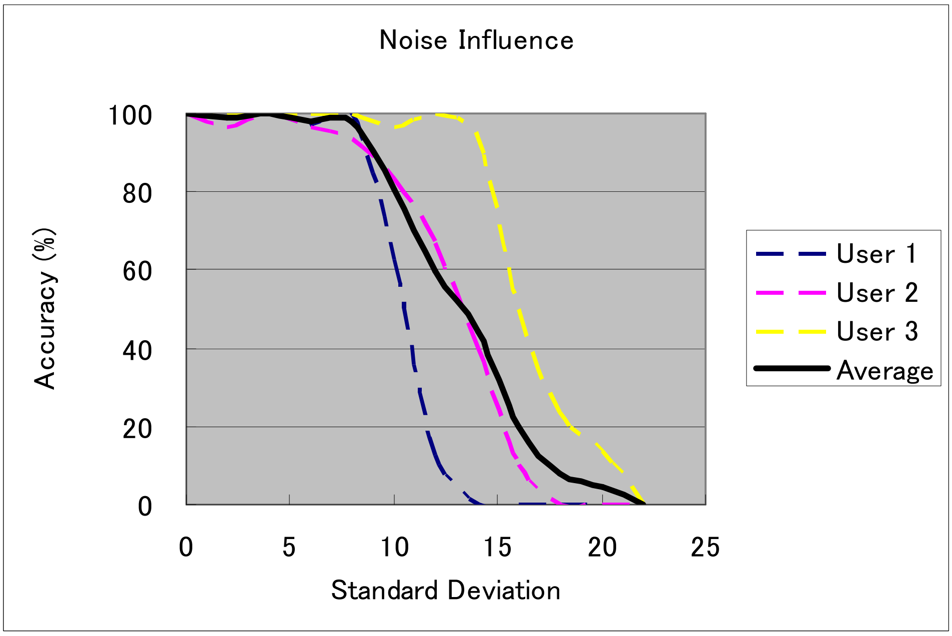



**FIGURE 12:** Noise Influence

Figure 12 shows the accuracy decrease as result of the adding of noise. The accuracy of user 1 decrease when standard deviation is around 9, the accuracy of user 2 decrease when standard deviation is around 10, and the accuracy of user 3 decrease when standard deviation is around 12. Between all users, it seems that user 3 is most robust against noise influence. As result, the experiment data of noise influence shows that noise gives small effect in blinking detection because Gabor filter eliminates it and the add of noise give small changes in the shape of eye arcs. Noise is detected as small point and it is automatically removed by Gabor filter as it only selects the horizontal line and ignores the others.

### 3.3 Comparison with Bio-potential based blinking detection method

The other approach of blinking detection is based on bio-potential analysis. Bio-potential instrument such as EMG (Electromyograph), EEG (ElectroEnchepalograph), and also EOG (Electrooculograph) have been used to detect the blinking. The most used blinking detection is based on EOG (by sticking the surface electrodes onto surrounding eyes in order to detect the eye muscle activity including eye blinking).


**FIGURE 13:** blinking detection by using NeuroSky blinking detector

Because EOG based on the activity of eye muscle, the accuracy of this approach is rely on how much force in the muscle. As the principle of works of muscle, it will shorter or expand based on the bio-potential which is given into this muscle. If the high bio-potential is given into the muscle, it has same meaning that the muscle is given with high power. This muscle will have high power to move something (to move the eyeball). Otherwise, if small bio-potential is given into muscle, the small power on muscle will move the eye slowly. In the bio-potential based blinking detection method, it will measure how much bio-potential is given into the eye muscle. When eye moves or blink with strong power, the EOG will detect the high bio-potential. Otherwise, when eye blink with less power, EOG will detect small bio-potential as result from blinking activity. This approach looks powerful to detect the blinking. Unfortunately, because the EOG signal is bellow 5 Hz, it means that the limit of blinking speed is 200ms. When the blinking speed is less than 200ms, the EOG will not be able to detect the blinking. Beside the EOG wave itself, the amplitude of bio-potential will also affect the accuracy. When bio-potential is very weak, the method will not be able to detect the blinking. Because of these reasons, the image analysis approach is more powerful than bio-potential approach in order to detect the blinking.

In order to measure the accuracy of bio-potential approach, we test the blinking of user by using NeuroSky Blinking detector as depicted in Figure 13. This tool will detect the signal from eye and determine the blinking. The user blinks 30 times. The detected blinking is 12 times. The result shows that there is error 60% (the accuracy is only 40%). This accuracy is caused that the NeuroSky instrument cannot detect the blinking if the blinking power is less.

### 3.4 Application of the proposed blink detection to computer input with human eyes only

One of the examples of application of the proposed method for computer input with human eyes only is introduced. Mouse is one of computer input devices which allows input the desired key into computer. Basically, eye-mouse has the same functions as those of general mouse. It, however, does work with human eyes only without any direct touch with computer. Eye-mouse is getting more important since it can be used for disable and handicapped persons as well as elderly persons who are not good at typing with keyboard to use computer. By using only one single camera, face image is captured. Based on the Viola-Jones method, face, mouth and eyes are detected and recognized. Based on the detected locations, two ends of left and right eyes as well as mouth and its surrounding region are extracted from the captured face image. Once these locations and their corresponding images are extracted, then they are tracked by the well-known template matching. Then their locations are mapped onto the 3D of real world coordinate system. 3 D head model is also created into the coordinate system. Eye gaze is estimated with the location of iris center and two ends of eyes. Absolute gaze is calculated from eye gaze itself and head location as well as attitude. From the eye gaze, location of which the user is looking at is estimated. To determine the selected location of key, eye blinking which is also detected by template matching is used. Also “fix one’s eyes on” can be used for determination of the selected key. Experimental results show that the success rate is 100% perfect when the user types the designated key on Microsoft WORD using screen keyboard. An example of computer screen image is shown in Figure 14. At the top right corner, program process monitor is situated while the screen keyboard is located at the top center. Below the screen keyboard, Microsoft provided WORD window is situated while users face and detected and tracked eye (left eye) and its surrounding images are located at bottom right corner. User can input any key in the screen keyboard by looking at the designated key for a while or blink the eye then selected key is determined. User can also confirm the location on the screen keyboard that he or she is looking at by the pointer location which is moved in accordance with the estimated gaze location. The locations of user's two eyes are also identified on the computer screen (in the user's face image situated at the bottom left corner). Thus user can input his or her designated key in the screen keyboard to the computer by human eye only. The proposed system also allows user's movement.

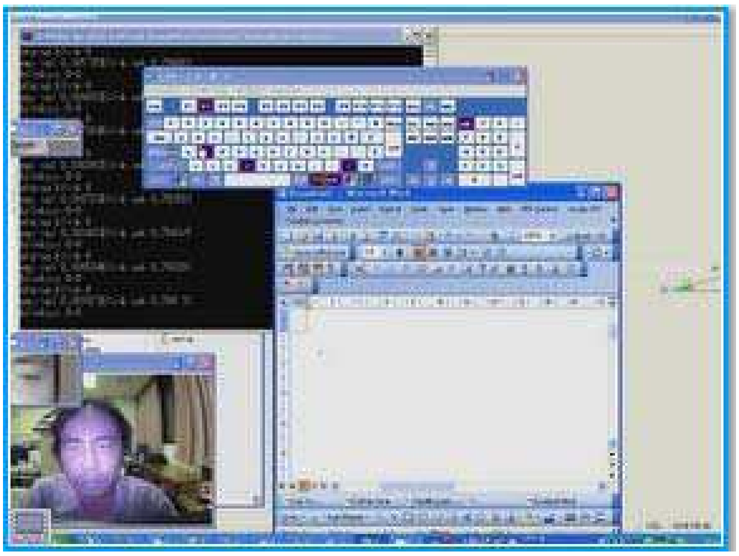


**FIGURE 14:** Example of the computer screen

.

## 4. CONCLUSIONS

The blinking detection based on Gabor filter has been successfully implemented. Gabor filter success to extract eye's arcs and estimate the blinking by using connected labeling method to measure the distance between arcs. The experiment data show that our proposed method superior than other methods against different users (perfectly detect blinking from three different users), varies eye shape, and has perfectly accuracy. It is also concluded that our proposed method robust against noise. The blinking detection method give better result than bio-potential based blinking detection since it is only detected by using camera and camera always give stable input than bio-potential voltage. By implement this method, the drowsiness detectors, HCI applications, and also Aid communications can be done with perfect accuracy and also robust against different users, varies eye shape, and noise.

## 5. REFERENCES

1. Robert Krupiński and Przemysław Mazurek. "Estimation of Eye Blinking Using Biopotentials Measurements for Computer Animation Applications". Lecture Notes in Computer Science, Vol 5337/2009, pp 302-310, 2009

2. Bin Abd Rani, M.S. bt. Mansor, W. "Detection of eye blinks from EEG signals for home lighting system activation". Proceeding of ISMA '09 6th International Symposium, pp 1-4, 2009

3. Ilkwon Park, Jung-Ho Ahn, and Hyeran Byun. “Efficient Measurement of Eye Blinking under Various Illumination Conditions for Drowsiness Detection Systems”. Proceedings of IEEE the 18th International Conference on Pattern Recognition, 2006

4. Kristen Grauman, Margrit Betke, James Gips, Gary R. Bradski. “Communication via Eye Blinks – Detection and Duration Analysis in Real Time”. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001

5. Zhu Hao, Qianwei Lei. "Vision-Based Interface: Using Face and Eye Blinking Tracking with Camera". iita, vol. 1, pp.306-310, 2008 Second International Symposium on Intelligent Information Technology Application, 2008

6. Yunqi, L., Meiling, Y., Xiaobing, S., Xiuxia, L., and Jiangfan, O. 2009. “Recognition of Eye States in Real Time Video”. Proceedings of the 2009 international Conference on Computer Engineering and Technology, Vol 01, pp 554-559, 2009

7. S. Sirohey, A. Rosenfeld and Z. Duric. “A method of detecting and tracking irises and eyelids in video". Pattern Recognition, vol. 35, num. 6, pp. 1389-1401, 2002.

8. T. Morris, P. Blenkhorn and F. Zaidi. “Blink detection for real-time eye tracking”. Journal of Network and Computer Applications, vol. 25, num. 2, pp. 129-143, April 2002.

9. Matjaz Divjak, Horst Bischof. “Eye blink based fatigue detection for prevention of Computer Vision Syndrome”. IAPR Conference on Machine Vision Applications, Tokyo, 2009.

10. Marc Lalonde, David Byrns, Langis Gagnon, Normand Teasdale, Denis Laurendeau. “Real-time eye blink detection with GPU-based SIFT tracking”. Fourth Canadian Conference on In Computer and Robot Vision, pp 481-487, 2007.

11. Michael Chau and Margrit Betke. “Real Time Eye Tracking and Blink Detection with USB Cameras”. Boston University Computer Science Technical Report No. 2005-12, May 2005.

12. Gang Pan, Lin Sun, Zhaohui Wu and Shihong Lao. “Eyeblink-based anti-spoofing in face recognition from a generic webcamera”. The 11th IEEE International Conference on Computer Vision (ICCV'07), Rio de Janeiro, Brazil, October 2007.

13. Akihiro Sugimoto, Akihiro Nakayama, Takashi Matsuyama. “Detecting a Gazing Region by Visual Direction and Stereo Cameras”. Proc. of the 16th International Conference on Pattern Recognition, 2002

14. Matjaž Divjak, Horst Bischof. “Real-time video-based eye blink analysis for detection of low blink-rate during computer use”. Institute for Computer Graphics and Vision Graz University of Technology Austria, 2008.

15. D. Gabor. “ Theory of Communication”. J. IEE, vol. 93, pp. 429-459, 1946.

16. J.G. Daugman. “Two-Dimensional Spectral Analysis of Cortical Receptive Field Profile”. Vision Research, vol. 20, pp. 847-856, 1980.

17. J.G. Daugman. “Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two- Dimensional Visual Cortical Filters”. J. Optical Soc. Amer., vol. 2, no. 7, pp. 1,160-1,169, 1985.

18. Tai Sing Lee. "Image Representation Using 2D Gabor Wavelets". IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 959-971, October, 1996.

19. http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/AV0405/SIKLOSSY/bars.html.

# Camera as Mouse and Keyboard for Handicap Person with Troubleshooting Ability, Recovery, and Complete Mouse Events

**Kohei Arai** arai@is.saga-u.ac.jp
*Information Science*
*Saga University*
*Saga, Japan*

**Ronny Mardiyanto** ronny_mardiyanto@yahoo.com
*Electrical Engineering*
*Institut Teknologi Sepuluh Nopember*
*Surabaya, Indonesia*

**Abstract**

Camera mouse has been widely used for handicap person to interact with computer. The utmost important of the use of camera mouse is must be able to replace all roles of typical mouse and keyboard. It must be able to provide all mouse click events and keyboard functions (include all shortcut keys) when used by handicap person. Also, the use of camera mouse must allow users do troubleshooting by themselves. Moreover, it must be able to eliminate neck fatigue effect when it is used during long period. In this paper, we propose camera mouse system with timer as left click event and blinking as right click event. Also, we modify original screen keyboard layout by add two additional buttons (button "drag/ drop" is used to do drag and drop of mouse events and another button is used to call task manager (for troubleshooting)) and change behavior of CTRL, ALT, SHIFT, and CAPS LOCK keys in order to provide shortcut keys of keyboard. Also, we develop recovery method which allows users break for a while and then come back again to eliminate neck fatigue effect. The experiments involving several users have been done in our laboratory. The results show that the use of our camera mouse able to allow users do typing, left and right click events, drag and drop events, and troubleshooting without hand. By using this system, handicap person can use computer more comfortable and reduce the eyes fatigue.

**Keywords:** Camera mouse, timer, blinking, mouse events, keyboard shortcuts, troubleshooting, fatigue.

## 1. INTRODUCTION

Recently, the numbers of handicapped persons who have limited in mobility have been rising. These may be caused by accident, congenital, diseases, and etc. These make person become cannot move their foot, hand, or other mobility organs. These limitations make them feel loose their life as they become dependent with others [9][10]. The rising numbers of handicapped person encourage the growing of computer input development to allow them interact with computer with their limitation to assist their daily life.

Computer input special design for handicapped person have been developed by many researchers. Computer input device, such as keyboard, mouse, joystick, and etc, have been widely used for computer interaction. Unfortunately, typical computer input device always requires hand's control. Controlling of cursor and selection of key must be done by using hand. To provide computer input for handicapped person, the possible solution is by utilizing camera as mouse and also keyboard. Mouse-keyboard utilizing web camera has been used to replace typical mouse and keyboard. It means that by utilizing web camera as mouse, it also can be used as keyboard by using screen keyboard. The method of controlling pointer in camera-based mouse-keyboard can be broadly classified into following categories: (1) Gaze-based (The pointer is controlled based on user's gaze) [2][3], (2) Face-based (The pointer follows center of face location) [1][6], (3) Head poses-based (The pointer is controlled based on head poses such as roll, tilt, and yaw) [4], (4) Hand gesture-based (The pointer is controlled by using hand gesture) [5], and (5) Laser based (Laser pointer mounted on user's glasses controls the pointer) [7].

Ref [2] developed eye mouse system based on user's gaze. After face is found and tracked, eye location is searched by projection of difference between left and right eye images. Output of this system is only left and right direction which used to control mouse pointer. No upward and downward directions are used. Ref [3] developed eye mouse system which user's gaze is obtained from pupil location. Also, blinking is used as left click mouse event. Ref [1] developed camera mouse using face detection and eye blink. Center position of face is detected by using Adaboost face detection method and tracked by using Lucas-Kanade method. This location is used as pointing value and blinking is used as left click mouse event. In ref [6], camera mouse is driven by 3D model based visual face tracking method. Human facial movement is decomposed into rigid movement (rotation and translation) and non rigid movement (facial expressions). Mouse cursor is controlled by face movement (x and y coordinates). Left click event is operated by opened mouth detection and right click event is operated by stretching mouth corner detection. Ref [4] developed camera mouse based on head poses. Head poses such as yaw, roll, and tilt are used to tune the mouse cursor. Rough position of mouse is estimated based on center of face. Right and left click mouse event are done based on roll angle. Ref [5] developed camera mouse based on hand gesture. Hand location is detected and tracked by using SIFT and CAMShift algorithm. Hand gesture is recognized by using PCA method. Mouse event is controlled by stretching the finger. Ref [7] developed unistroke keyboard and mouse by using image of a keyboard (is made from paper) and laser pointer mounted on user's glass. The keyboard image is monitored by a camera and illuminated key locations are recognized. According to the type of the click mouse event, it can be classified into (1) Left click event only [1] [3] and (2) Left and Right Click events [4] [5] [6]. There are several methods such as timer, blinking, roll movement of head, stretching of finger, and mouth condition are used to control mouse events. Refs [1] and [3] used blinking as left click event. Refs [4], [5], and [6] used roll movement of face, stretching of finger, and state of mouth to control right and left click mouse event. The comparison between click mouse events of each method is shown in table.1.

| **Ref** | **Click Events** | |
|---|---|---|
| | **Left Click** | **Right Click** |
| [1] [3] | Blinking | None |
| [2] [7] | None | None |
| [4] | Roll movement of face | Roll movement of face |
| [5] | Stretched finger | Stretched finger |
| [6] | Opened mouth | Stretched mouth |

**TABLE 1:** Comparison between existing camera mice according to their click events method

The existing camera mice only consider how to select and choose the desire point. There is no camera mouse system which provides complete click events (right and left click events), allows shortcut keys of keyboard, and also allows user do troubleshooting when not-responding program happens. Moreover, the uses of camera mouse during long period make user's eyes become dry moreover give neck fatigue effect. For the aforementioned reason, we propose camera-based mouse-keyboard system with troubleshooting ability, recovery, and complete click mouse events by using timer and blinking.

The objectives of this proposed system is to overcome click mouse events, shortcut keys of keyboard, troubleshooting, recovery, eyes dryness, and neck fatigues effect problems of previous camera mice system. Mostly, the existing camera mice only provide left click mouse event. Only Refs [5] and [6] provide both of left and right click mouse events. It utilize stretched finger and opened mouth to control left and right click mouse events. Because the use of camera mouse as keyboard always involve screen keyboard, it must allow all keyboard shortcut keys. Unfortunately, the original screen keyboard didn't provide hold function of CTRL, ALT, SHIFT, and CAPS LOCK keys and causes system eliminate shortcut keys such as CTRL+C as copy, CTRL+V as paste, and etc. When computer freezes, the typical solution is by press CTRL + ALT + DEL keys simultaneously to show task manager. Unfortunately, the use of camera mouse for handicapped person eliminates this shortcut because user cannot press CTRL + ALT + DEL keys simultaneously by hand. Also, when users want to break, the system should be able to recover and able works again when user comes back to front of computer. The use of camera mouse during long period makes user's neck and eyes become tired. Also the use of blinking as left click event makes user's eye become dry.

In this paper, we propose camera mouse system combining timer and blinking as right and click mouse events. Left click event is controlled by timer and right click event is controlled by blinking. Our proposed system is different with other methods because it provides complete click mouse events. This method allows users do all of click mouse events. Moreover, we modify windows screen keyboard layout by add drag/ drop, task manager buttons, and change the behavior of CTRL, ALT, SHIFT, and CAPS LOCK keys in order to allow shortcut key of keyboard. The drag/ drop button allows users drag and drop the desired object. Instance, when user want to move the window of program, users click "drag/ drop" button first and continue with left click. This first click will drag the object until the second click happens. This makes window program able to be moved to another place and drop will be automatically selected when the second left click happens. Also, task manager button allows users do troubleshooting by themselves. When system freezes, user can end task the active program by call the task manager. It will replace the use of "CTRL – ALT – DEL" combination keys. Moreover, the visibility of user face is monitored by using three tracked points. When users face is invisible, system will automatically reboot and back to the beginning process. This kind of recovery process will eliminate neck fatigue effect and eye's dryness.

This paper is organized as follows: section 2 describes our proposed system involving hardware configuration, modified screen keyboard, mouse click events, and recovery method. Section 3 describes our experimental results involving performance camera mouse when used for typing, blinking detection accuracy, and testing of recovery. Section 4 describes the conclusions.

## 2. PROPOSAL

The problem of existing camera mouse is how the use of camera mouse can replace all functions of typical mouse and keyboard. As we now that typical mouse has two click events (left and right click), left click is used for selection and right click is used to show the additional menus. The system should provide drag/ drop events without burden the user. Because the use of camera mouse automatically leaves the use of hand, it must be able to play keyboard's role as one of computer input device. Beside it can be used for typing, it must be able to do troubleshooting. When computer become freeze, by press "CTRL - ALT - DEL" user can end task the not-responding program. The use of camera mouse must cover whole functions of typical mouse and keyboard. Pointing, clicking events, typing, and troubleshooting must be done by camera mouse only. Moreover, the use of camera mouse should not burden the user (make user's neck fatigue).
When we use camera mouse as computer input, it should be considered about how the user do typing, drag/ drop event, troubleshooting, right and left click event, and others. To provide typing function, almost camera mouse use screen keyboard software. The screen keyboard software with keys button layout allows user type by clicking the key button. Actually, screen keyboard

software can replace the typical keyboard's function. Unfortunately, screen keyboard program doesn't allow user types "CTRL - ALT -DEL" simultaneously. It means that the use of screen keyboard eliminate troubleshooting keys because user cannot open the task manager program by pressing "CTRL - ALT -DEL" simultaneously.

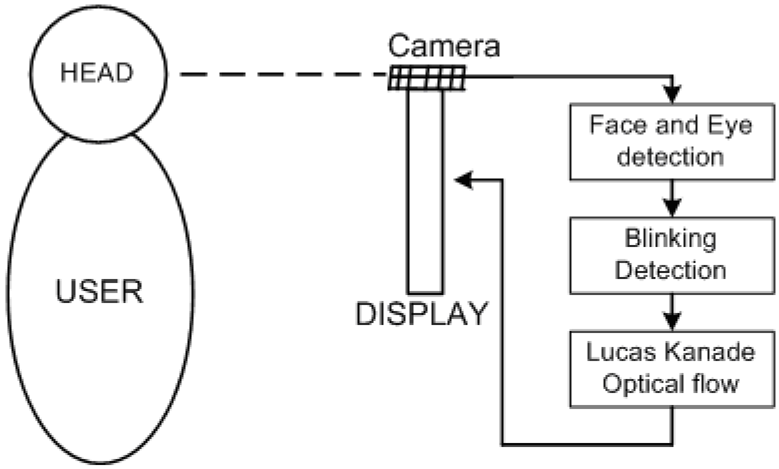



**FIGURE 1:** Block diagram of proposed system.

This paper proposes camera mouse with contribution involving modification of screen keyboard layout with task manager and drag/ drop buttons, left click is done by timer, right click is done by blinking, and the recovery method allowing user go away and go back from/ to front of camera in order to avoid neck fatigue and dryness of eye. By adding of task manager button, it allows user does the troubleshooting by him/ her self. Also, after drag/ drop button is clicked, once user click the object, this click is recognized as drag event and next click is recognized as drop event. The use of timer and blinking make camera mouse system acting as typical mouse (with left and right click events). To update the cursor position, once the face location is found by Adaboost based on Haar-like feature method (proposed by Viola and Jones), Lucas-Kanade Optical flow track this position and use to update cursor position. Last, we use normal flow to detect blinking. The block diagram of proposed system is shown in Figure 1.

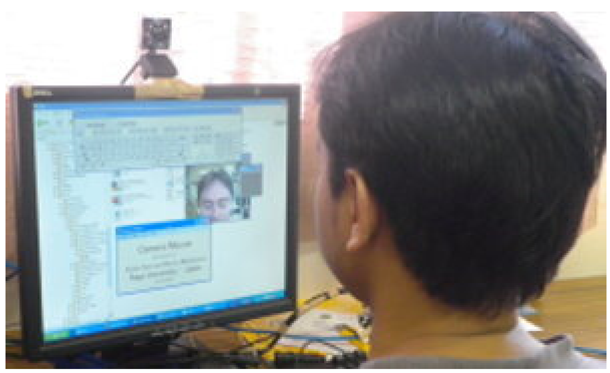


**FIGURE 2:** Proposed Camera Mouse System.

### 2.1 Hardware Configuration

We develop our system using Visual Studio 2005 and OpenCv Library downloaded as free on their website. Optiplex 755 dell computer with Core 2 Quad 2.66 GHz of CPU and 2G RAM is used. Web camera NetCowboy DC-NCR 131 with 1.3 M pixels resolution is used as main source. Web camera is placed on top of display. The distance between user and camera is 30 cm. Even though we used high speed computer, we have tested that our system works with ASUS EEEPC 1005HE netbook (Intel Atom N280 processor with 1.66 GHz, 1GB RAM, and built-in1.3 M pixels web camera). Also the use of our system is not limited for Infra Red camera but also invisible camera. It means that all type web cameras with minimum resolution 1.3 M pixels also can be used. The proposed camera mouse is shown in Figure 2.

## 2.2 Modified Screen Keyboard

The objective of this modification is to allow users do drag/ drop click events and troubleshooting. The existing camera mouse systems do drag/ drop events by utilizing blinking. The closed eye can be signed as drag event and when opened eye is signed as drop event. The weakness of this method is that drag/ drop events are always rely on eye while eye condition will become tired/dry during long use duration. When eye is used only in a few minute, it will not make eye become dry. Unfortunately, the use of camera mouse is usually in long duration depending on necessity. The modified screen keyboard layout is shown in Figure 3.


**FIGURE 3:** Modified Screen Keyboard

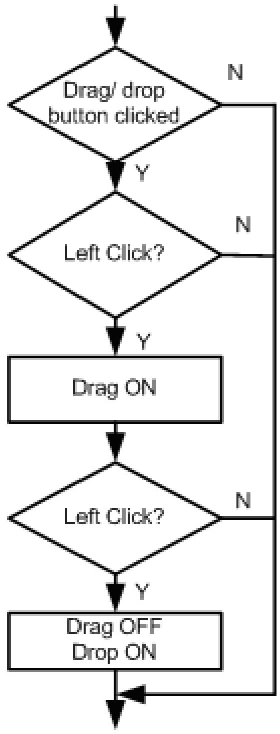



**FIGURE 4:** Process flow of drag and drop events.

We modify the screen keyboard layout by adding two buttons. The first button is task manager button assisting user when not-responding program happens. It means that the troubleshooting of computer which usually do by pressing "CTRL + ALT + DEL" keys, now is replaced with this button. When user uses the camera mouse and not-responding program appears, user can click this button to end task the not-responding program. The second button is drag/ drop button for assisting user do drag and drop the object. Instance, when user want to move the window position of running program, user can click this drag/ drop button and continue with click to the window and then the last click is automatically signed as drop event. The adding of this button replaces function of drag and drop mouse click events. The process flow of drag and drop events are shown in Figure 4.

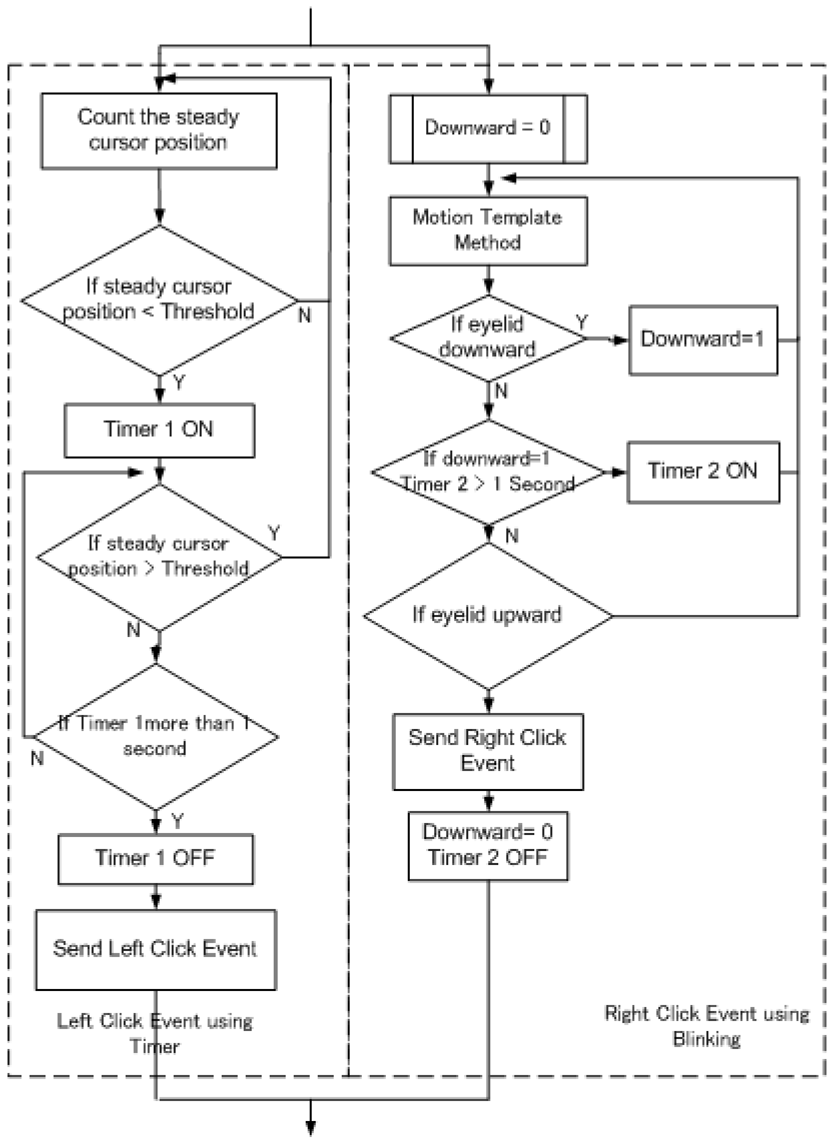



**FIGURE 5:** Click event process flow.

## 2.3 Mouse Click Events

To interact with computer, mouse has been used by user to give input on graphical user interface by using clicking. The stored procedure can be controlled by using mouse. It means that mouse assist user does interaction with computer by clicking the button or other graphical user interface tool. The typical mouse usually has two buttons: (1) left click button (performs a first action on object selection) and (2) right click button (performs a secondary action). After mouse cursor is controlled, the decision is done by these click events. Instance, when mouse is used for typing using screen keyboard software, user select the desired key by clicking the key. Typically this clicking is done by using left click. Also, when user wants to show the properties menu of icon, user does right click on this icon. Both of left and right clicks can be used simultaneously and it make user become easy to interact with computer. In this paper, we propose mouse click events using timer and blinking. When camera mouse is used to replace the role of typical mouse, it also must cover the complete mouse click events. Left and right click must be provided by camera mouse system. Our system use timer as left click event. By using timer, left click will automatically click when cursor position stops during specific duration. This method is easy to use than other method such as blinking, head rotation, opened mouth, and etc because there is no physiological effort is required. According to that left click event is more often used than right click event, the method must not burden the user physiology. We cannot use blinking method as left click because the use of blinking during long period causing the user's eye becomes dry. We use blinking as right click event. The click events process flow is shown in Figure 5.

The both click events run simultaneously. In left click event, system counts the steady of cursor position. If the steady of cursor position less than threshold, timer 1 is activated. However, when steady of cursor more than threshold, timer 1 is resetting. If timer 1 more than 1 second, system send left click event while timer 1 is turned OFF. The steady of cursor position is calculated as follow,

$$\mu(x, y) = \sum_{i,j} \frac{P(x_i, y_j)}{N}, \quad (1)$$

$$Steady = \sum_{N} \left| P_N(x) - \mu_N(x) \right| + \left| P_N(y) - \mu_N(y) \right|, \quad (2)$$

where P(x, y) is cursor location and N is number of locations. The left click event by using timer is shown in Figure 5 (left).

To provide right click event, we use blinking. Blinking is detected by using motion template method [8]. This method obtains magnitude and angle of total motion on image. After eye image is cropped, motion template method estimates the motion. We detect blinking if magnitude of motion on image is more than threshold and the first angle is 270º (eye become closed) and the opened eye is signed by angle is 90º. The blinking detection method is shown in Figure 5 (right). First, we detect the closed eye by using motion template method. When eye become closed, timer is turned ON. If eye become open when timer 2 bellow than 1 second, it will not be recognized as blinking. Otherwise, if eye become open and timer 2 exceed 1 second, it will be recognized as blinking. If blinking happens, system send right click event.

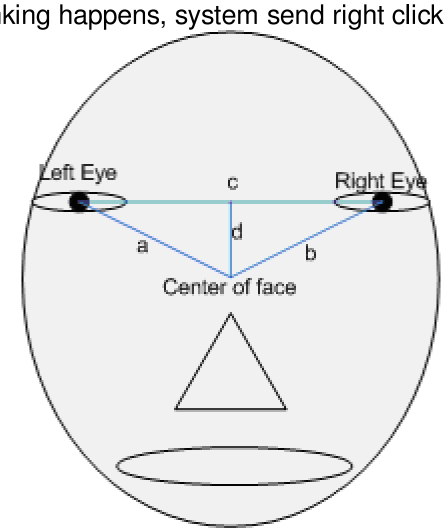



**FIGURE 6:** Detection of face dissipation caused by head movement.

### 2.4 Recovery Method

To avoid neck fatigue and dryness of eyes, we develop recovery method. This method prevents error when user's face is invisible. When user's neck becomes tired, user may take a rest by going away from camera. Because the face is invisible, error may happen due to this situation.

This kind of error can be avoided by resetting all variables and return to the beginning. The action of dissipation of user's face can be divided into: (1) out of region (when user shifts the position until out of the camera boundary) and (2) dissipation that caused by head movement (yaw and tilt). When user shifts the face until out of region, the all variables are reset and return to first process. However, disappearing of user's face that caused by head movement is detected by measuring the reasonable distance between center face and eyes. If the distances of face and eyes are more than threshold, then all variable is reset.

Figure 6 shows the detection of dissipation of face which caused by head movement. The over yaw movement happens when |a-b|> threshold. The over tilt movement happens when d > threshold. When the over head movements are detected, system will reset all variables and return to beginning.

## 3. EXPERIMENTAL RESULTS

To test the camera mouse performance, we conducted several experiments: measurement of blinking accuracy, testing of recovery (measuring the camera mouse performance when user go from front of camera and then come back again to front of camera), and testing of typing performance (use the camera mouse to type on MS Word). The detail experiments are explain bellow,

### 3.1 Blinking Accuracy

The objective of this experiment is to measure the accuracy of blink detection method. We measure the blinking accuracy by compare between manual detection and our blinking detection method. This experiment is done by users blink their eye 20 times. The accuracy of blinking detection is 90%.

### 3.2 Recovery Testing

The objective of this experiment is to know how our system can recognizes invisible user's face and also recover when user comes back again to front of camera. We counted the time of recovery process for ten times. The average time for recovery process is 8.7 seconds.

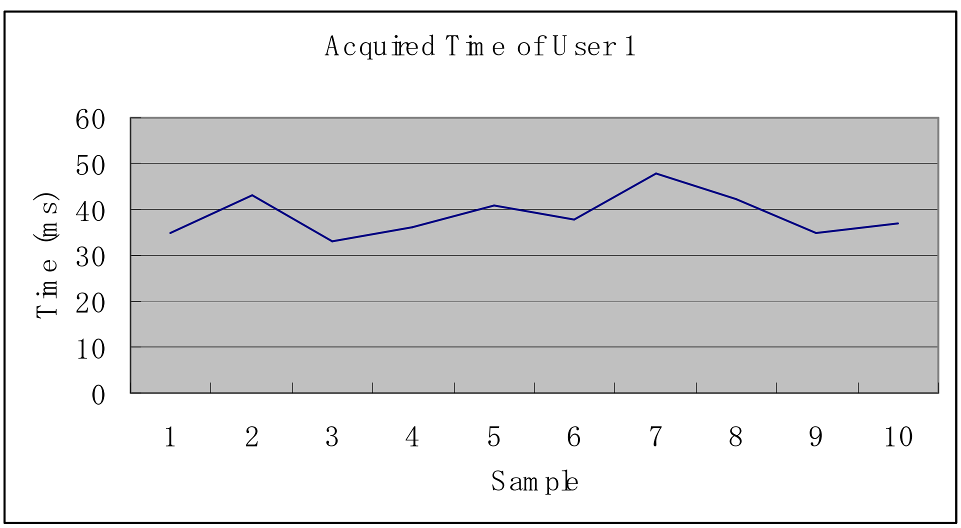



**FIGURE 7:** Acquired Time of User 1. This figure shows that the acquired time is stable and relative fast.

### 3.3 Typing Performance

To test the camera mouse usability, we conducted experiment with user typing on MS Word by using our camera mouse. This experiment involves four different users with each of them has different nationality. User sit down in front of computer with web camera mounted on the top of display and types “Saga University” on MS Word using our camera mouse. User types “Saga University” until ten times while the consumption times are recorded. The recorded times are shown in Figure 7, Figure 8, Figure 9, and Figure 10. The first user is expert user. He is Indonesian. He often used our software for typing. He can fast and easy control the cursor of our camera mouse. The acquired time of first user is shown in Figure 7.

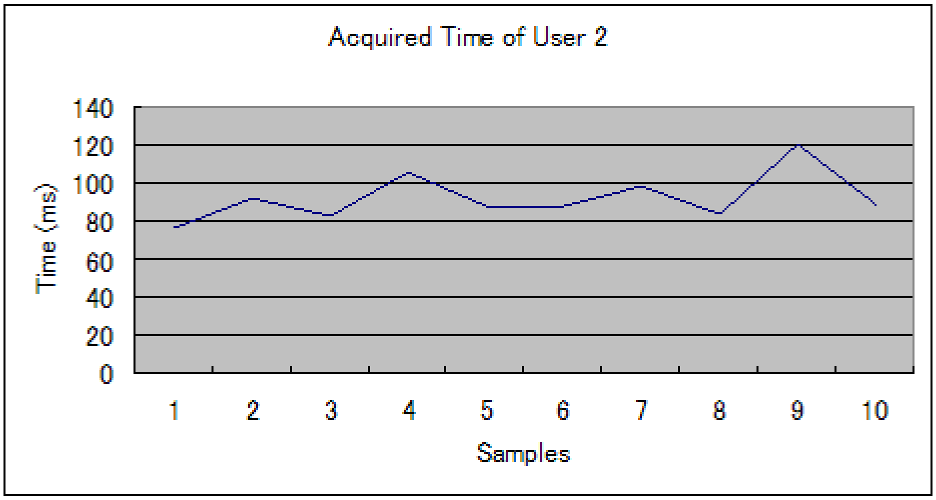



**FIGURE 8:** Acquired Time of User 2. This figure shows that the acquired time rise at last sample. At last sample show that user become tired on his head.

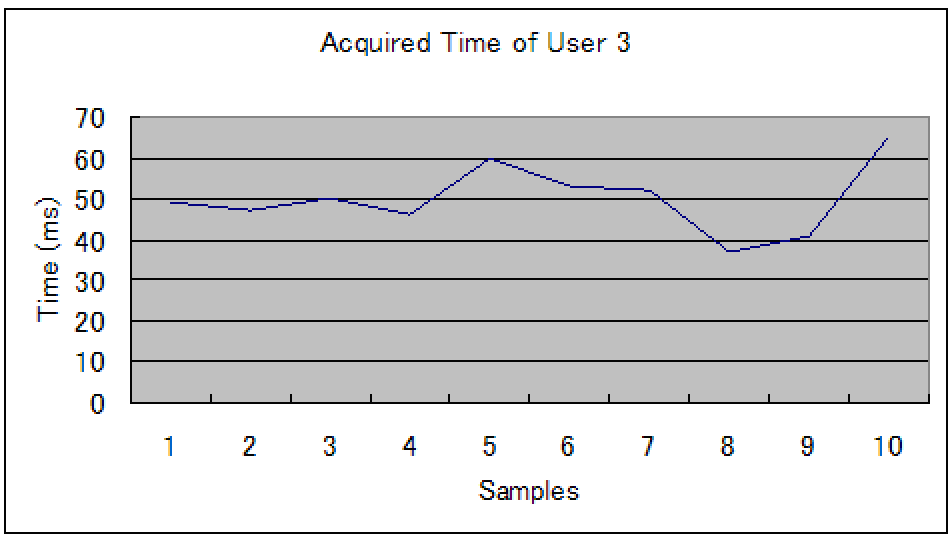



**FIGURE 9:** Acquired Time of User 3. This figure shows that between first samples until 7th sample, user able stable to type on MS Word. At 8th and 9th samples, user able type faster than previous because he is lucky and no mistake happen. Unfortunately, at 10th sample user become tired and he made several mistake and caused longer time to type.

The second user is beginner user. He is Sri Lankan. This experiment is the first time for him using our camera mouse. Before the experiment, he never used it before. Because of this, we explained how to use this and let him practiced for one minute. After one minute practiced, he began to type “Saga University” on MS Word by using our camera mouse. The recorded consumption time is shown in Figure 2. In this figure, we can see that the time will increase when user uses our camera mouse during long period. It is caused by he feel that his neck becomes tired. His neck will become tired because he never used his neck for controlling something before. If he uses it everyday then he will not feel tired again. The next user is also beginner user. He is

Indonesian. He never used our camera mouse before. After did practice for one minute, he began to type “Saga University” using our camera mouse. The acquired time is shown in Figure 3. The last user is also beginner user and never used it before. He is Bangladeshi.

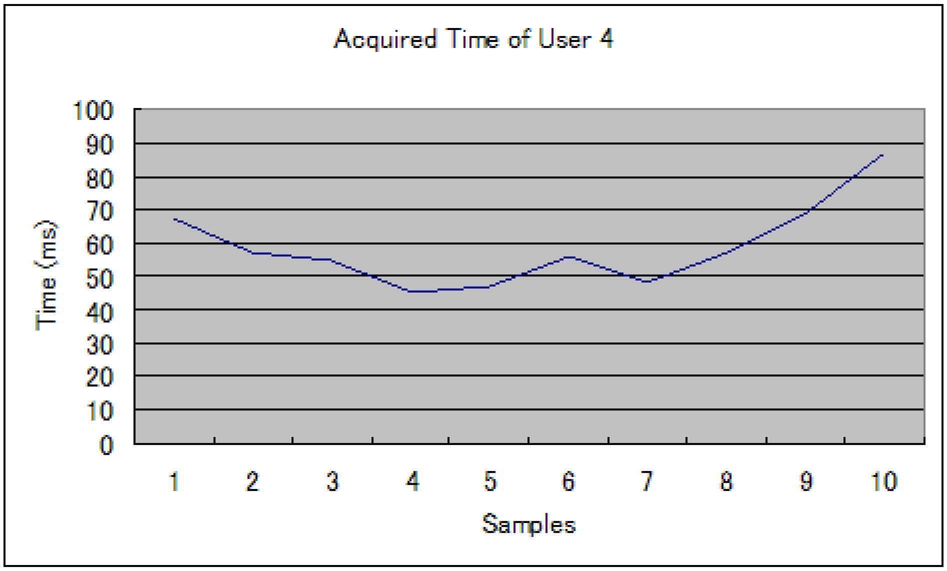



**FIGURE 10:** Acquired Time of User 4. This figure shows that the acquired time is decrease at four of first samples. It is caused that user need to adapt with our camera mouse. After 4th sample shows that it becomes stable because user becomes familiar with it. The two of last sample show that acquired time is increase. It is caused by user neck become tired.

### 3.4 Works with Others Windows Application

We used our camera mouse to control windows application such as MS Word (For typing), Soliter Game (Using drag and drop functions), selecting file from windows explorer, playing music by using windows media player, setting properties of file by using blinking as right click event, setting windows volume, and etc. User used our camera mouse system which acts as mouse and keyboard simultaneously.

## 4. CONSLUSIONS

The camera mouse with ability to replace all mouse and keyboard functions has been successfully implemented. By using timer and blinking, users successfully do left and right click events, drag and drop events, and also troubleshooting. Combination between timer and blinking, as left and right click events, has advantage in reducing fatigue because timer can be used more often than blinking, so timer is suitable to replace left click meanwhile blinking is used to replace right click event. The added button allows user did troubleshooting by clicking the task manager button from our modified screen keyboard layout. Another added button also allows user use shortcut combination keys such as “CTRL+A”, “CTRL+C”, and etc. By implement this system, handicapped persons will able to use computer completely.

## 5. REFERENCES

1. Zhu Hao, Qianwei Lei, "Vision-Based Interface: Using Face and Eye Blinking Tracking with Camera," iita, vol. 1, pp.306-310, 2008 Second International Symposium on Intelligent Information Technology Application, 2008.

2. John J. Magee, Matthew R. Scott, Benjamin N. Waber, Margrit Betke, "EyeKeys: A Real-Time Vision Interface Based on Gaze Detection from a Low-Grade Video Camera," cvprw,

vol. 10, pp.159, 2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'04) Volume 10, 2004.

3. ChangZheng Li, Chung-Kyue Kim, Jong-Seung Park, "The Indirect Keyboard Control System by Using the Gaze Tracing Based on Haar Classifier in OpenCV," ifita, vol. 2, pp.362-366, 2009 International Forum on Information Technology and Applications, 2009.

4. Yun Fu, Thomas S. Huang, "hMouse: Head Tracking Driven Virtual Computer Mouse," wacv, pp.30, Eighth IEEE Workshop on Applications of Computer Vision (WACV'07), 2007

5. Hailing Zhou, Lijun Xie, Xuliang Fang, "Visual Mouse: SIFT Detection and PCA Recognition," cisw, pp.263-266, 2007 International Conference on Computational Intelligence and Security Workshops (CISW 2007), 2007.

6. Jilin Tu, Thomas Huang, Hai Tao, "Face as Mouse Through Visual Face Tracking," crv, pp.339-346, The 2nd Canadian Conference on Computer and Robot Vision (CRV'05), 2005.

7. M.E. Erdem, I.A. Erdem, V. Atalay, A.E. Cetin, "Computer vision based unistroke keyboard system and mouse for the handicapped," icme, vol. 2, pp.765-768, 2003 International Conference on Multimedia and Expo - Volume 2 (ICME '03), 2003

8. Gary Bradski, Andrian Kaebler: ”Learning Computer Vision with the OpenCV Library”, O’REILLY, 341-348, (2008).

9. National Institute of Neurological Disorders and Stroke, http://www.ninds.nih.gov/disorders/sci/detail_sci.htm

10. Christopher & Dana Reeve Foundation, http://www.christopherreeve.org

# A Proposed Web Accessibility Framework for the Arab Disabled

**Mohammed Abo El-Soud** mohamed_hossieny@yahoo.com
*Faculty of Computers & Information Sciences /*
*Computer Science Department*
*Mansoura University ,*
*Mansoura,*35516 *, Egypt*

**Samaa M. Shohieb** Sm.shohieb@yahoo.com
*Faculty of Computers & Information Sciences/*
*Information System Department*
*Mansoura University ,*
*Mansoura,*35516 *, Egypt*

**Abstract**

E-learning is a useful tool that has contributed in facilitating education for people around the globe. There is a need, however, for making e-learning available to disabled people.
This paper presents a Web accessibility framework which offers the ease of the Web accessing for the Arab disabled users and facilitates their lifelong learning. The basic objective of this framework is the support of the equal rights of Arab disabled people for their access to the education and training.

**Key Words :** Arabic Moon code, Arabic Sign Language, Deaf, Deaf-blind, E-learning Interactivity, Moon code, Web accessibility , Web framework , Web System, WWW.

## 1. INTRODUCTION

The Web is providing unprecedented access to information and interaction for people with disabilities.
Web accessibility basically means that people with disabilities can use the Web. More specifically, Web accessibility means that people with disabilities can perceive, understand, navigate, and interact with the Web [1]. However, Web accessibility is not a reality throughout the Web. The problem is that most websites have accessibility barriers that make it difficult or impossible for many people with disabilities to use them. And most Web software tools are not sufficiently accessible to people with disabilities, making it difficult or impossible for them to contribute to the Web. This means that efforts are needed to build a system which makes the Web accessible especially for the disables.
There is a growing, worldwide recognition that users with disabilities have the same right as others to access information technologies [2].
This paper proposes a Web accessibility framework for the main types of the disabilities (deaf-blind/ deaf/ blind) to facilitate the Web accessing for the Arab disabled. The proposed framework enables them to make use of the different websites with low efforts, low time and very low costs. The proposed framework depends on extracting the website content via reading and analyzing the meta-language of any Web page. Then the extracted content is presented in a format that fits the disabled user.
The paper is structured as follows. The next section presents the previous work that has been built for each type of the disabilities. Section 3 presents the proposed Web accessibility

framework. Section 4 presents the proposed Arabic Moon code for the deaf-blindness. Finally Section 5 summarizes the paper, and outlines possible avenue for future improvement.

## 2. RELATED WORK

Surveyed work done in [3] estimated the number of people with certain disabilities and “access” to the Internet. What “access” means is ambiguous, though, by the researchers’ own admission: It could simply mean a computer exists in the home or workplace that can be connected to the Internet or it could refer to active Internet use by the person in question. Even estimated about 43.3% of the world population have some kind of a disability By contrast, in the same survey, 56.7% of non-disabled people have Internet access. The disparity is considerable [4].

The survey is categorized as follows:

- Deafness Related Work: It contains the available accessibility systems that deal with the deaf persons.
- Deaf-Blindness Related Work: It contains the available accessibility systems that deal with the deaf-blind persons.
- Blindness Related Work: It contains the available accessibility systems that deal with the blind persons

### 2.1 Deafness Related Work

There are four categories of that type of survey based on the use of this system and the user category that interact with it.

•Interactive sign language learning systems: Systems in this category are focused on teaching the deaf students in different countries.

•Content producing systems: Systems in this category is used to provide conversion tools which are used for constructing, storing and maintenance for educational material.

•Text-To-Sign Browsers: A system in this category automatically converts the Web page content to SL. Research at that category is rare and still at the beginning.

•Online Sign Languages Dictionaries: Systems in this category play a very important rule in learning the SL for different languages. All the available SL dictionary systems aren't dedicated for the Arab deaf users' .Moreover; they are video based systems so they have low reliability.

#### *2.1.1 Interactive Sign Language Learning Systems*

Khwaldeh et al in [5] proposed a centralized based learning system, which aims to facilitate teaching and learning for both teachers of the deaf and deaf people. This system enables teachers and deaf to interact with each other. But this system is still limited in its use.

In [6] Gennari et al presented LOgic-based e-tool for DEaf children (LODE) that aims at stimulating deaf children to globally reason on narratives written in Italian. Thus LODE presents children with e-stories and apt exercises that stimulate them to analyze the temporal relations between events, and to produce new relations consistent with the story .But this system doesn't support the Arabic language as it supports only the Italian language .

Kyun Ng et al in [7] proposed an E-Learning framework that creates a common platform for both normal and disabled students which will share the same influential of their academic achievement. The deaf students can communicate with instructor and other students by messaging over the chat-room system .But this system need training on using it and it doesn't support learning the SL in general.

To encourage the deaf children to learn American Sign Language (ASL), Shirali-Shahreza et al [8] proposed a system which is implemented using PHP scripting language [9]. When a deaf person wants to enter a website which is created for deaf persons, a word is shown as a movie using SL. The user should recognize the word and select it from a list. If the user understands the SL and recognizes the word, he/she can enter the website. This system isn't used for e-learning as it could be used for Web browsing.

Another system is proposed in [10] which is an interactive program to teach ASL for K-3 mathematics by 3D animation. But this system is limited for a special type of courses and ages.
Straetz1 et al in [11] proposed a Learning Management System (LMS) which offers German SL videos in correspondence to every text in the learning environment. But it doesn't support Arabic Sign Language (ArSL). And Drigas et al in [12] presented a similar LMS but for Greek SL. The systems are designed notably for deaf adults who want to maintain and improve their mathematical and reading/writing skills. These systems require large bandwidth for downloading videos. So the reliability of these systems is low.
Ohene-Djan et al [13] proposed a system, Kids Sign Online (KSO) system, which is designed to teach British Sign Language (BSL) in tandem with English to deaf children. But this system supports a specific language BSL and doesn't support ArSL.
Stewart et al in [14] proposed a library that is available on the ASL browser website for teaching the ASL. For each word, the “ASL Browser” site has a movie shows a person saying the word using ASL .The movies are in Quicktime [15] format. Although this system supports a small sized movies but its reliability is still low.
There are still limitations in developing e-learning applications which use the ArSL for teaching deaf Arab students, and the already existent rare ones [16, 17] are missing interactivity between the user and the system.

***2.1.2 Content Producing System***
In [18] Webducation Software Planungs- und EntwicklungsgmbH carry out a project in close cooperation with the Austrian Association for hearing impaired and deaf people (WITAF). The objective of the project is to give possibility for deaf to generate and publish contents by themselves on the learning platform
Efthimiou et al in [19] proposed a platform environment that allows the development of various educational applications accessible by deaf users for Greek Sign Language (GSL).

***2.1.3. Text-To-Sign Browsers***
In [20] Boldyreff et al proposed a text-to-sign browsers for users of BSL. It is currently limited in their use. One of the major problems is that BSL and ASL do not translate word for word into English or vice-versa, as they have distinct grammars of their own, and therefore it would be difficult to translate a site directly .This type of systems isn't available for ArSL.

***2.1.4 Online SL Dictionaries***
Dasgupta et al in [21] proposed a cross platform multilingual multimedia Indian Sign Language (ISL) dictionary building tool. However; this system doesn't provide the ArSL .Troelsgård & Kristoffersen in [22] proposed a Danish Sign Language (DaSL) dictionary. For signers who have Danish Sign Language as their first language, the dictionary will provide information about Danish Sign Language such as synonyms and variants. This dictionary is a monolingual dictionary.
Vettori et al in [23] presented Electronic Bilingual Dictionary of Italian Sign Language (ISL) and Italian but it again doesn't support ArSL.
There is also another e-LIS dictionary proposed in [24] which is the Web bidirectional dictionary for Italian sign language-Italian.
Mohandes in [25] proposes a system that translates Arabic text to Arabic sign language. Words that correspond to signs from the Arabic sign language dictionary calls a pre-recorded video clips showing the sign. If the word does not have a corresponding sign in the sign language dictionary, it is finger spelled .But , this system depends on the videos and ,videos need a large bandwidth for downloading.
Suzuki et al in [26] proposed a Japanese and American Sign Language Dictionary System for Japanese and English users. This dictionary again doesn't support ArSL.
The Online Sign Languages Dictionaries [27-37] constitute very important educational tools for the e-learning and training of Sign Languages. To this direction many online dictionaries have been developed for different Sign Languages. The majority includes a large number of signs and targets at signers as well as at students that learn a Sign Language as second language. Each sign is accompanied from the material and includes, on one side one translational

equivalent and the other side synonyms and antonyms in the Sign Language. All that dictionaries don't support the Arabic language.

### 2.2 Deaf-Blindness Related Work

It was the advent of the personal computer with Braille or magnified visual output that opened up opportunities for a significant increase in access to information for deafblind people [38]. Software for producing large characters on the monitor is relatively inexpensive, but Braille displays have remained expensive.

Since there is a shortage of skilled transcribers, computer systems are often used to translate text to contracted Braille which is then output on a special embosser [39].

The algorithms for this translation are not simple since the rules governing the use of contractions depend on pronunciation and meaning. For example, there is a contraction for 'mother' which can be used as part of a longer word as long as it does not bridge a syllable boundary as in 'chemotherapy'. [39]

The DOS text-based operating system is easier for many deafblind people than ones, such as Windows, which use a graphical user interface. However keeping to DOS restricts the choice of software in that most new software is written for the Windows environment [39].But it isn't recognizable to use the DOS in the existence of the Web and the Web pages .

Over the years a number of systems have been developed to emulate finger spelling since many deafblind people do not read Braille. Hiroshi & Chikamune in [40] proposed a communication device that could help the deaf-blind communicate with others who do not know the Braille or sign language. Although these devices work well in a laboratory, there have been problems in making them generally available at affordable prices.

Fu & Ho in [41] described a finger language recognition system for handicapped aphasiacs, who are able to express their intents only by using ‘finger language’. Finger language is different from sign language in the sense that it is composed of simple hand gestures, each representing a predefined meaning.

The increasing use of graphics in printed books gives problems. Although many diagrams can be converted to an embossed form, the process of reading by touch means that a diagram has to be tactually scanned and mental image built of the whole diagram [39].This is the opposite process to visual reading where one looks at the overall picture and then reads the detail.

Years ago NASA [39, 42] had a problem with communicating with astronauts during lift-off. The problem was of information overload using visual and auditory communication. Therefore they investigated the use of tactual communication; the project failed, but the research formed the basis of a reading aid for blind and deafblind persons

Systems to recognize printed characters have been developed for inputting text to computers. Such systems have immediate application for deafblind persons since the information can be output in Braille [42].

#### 2.2.1 Moon Code

Moon allows people who are blind or partially sighted to read by touch. It is a code of raised shapes and takes its name from its blind English inventor [46].

There are many advantages for using Moon code [42, 45] which are:

Moon provides an "active" reading method for people who cannot access print - listening to audio books etc is passive

Self study courses in Grade 1 and Grade 2 Moon are available from RNIB, enabling a would-be learner to make a start even if a teacher is not available

Moon is larger and the characters are more "open" than Braille, so easier to feel and decipher

Moon requires a considerably less acute sense of touch than Braille, so can sometimes help readers with diabetes whose finger sensitivity is reduced

Some children and adults with learning and/or physical difficulties in addition to their visual impairments acquire some literacy through Moon, where Braille would be impossible.

Depending on that Moon is not widely known, there are some disadvantages [44-47] of it.

The choice of Moon books available is very limited at present.

There are currently no Moon magazines available, except for deafblind readers.

There is no portable, mechanical device for writing Moon, which there is for Braille.

The number of Moon readers is low and declining.
Whereas a "soft Braille" display can be linked to a computer in order to know what is on the screen, the Moon equivalent is not available, which is a particular disadvantage to those who cannot use a speech package.
All of the available dedicated systems for deaf-blindness have some problems such as high cost and have low reliability. Overall, they can't deal with the Arabic language.

### 2.3 Blindness Related Work

There are many screen readers and screen magnifiers to help the visual impaired as in [48; 49]. But all the available dedicated systems for blind users are very weak with the Arabic language and they aren't specified for the Arabic Web users. Consequently; researcher tried to solve this problem in a simple way in the proposed Web accessibility framework as an analysis and design stages.

## 3. THE PROPOSED WEB ACCESSIBILITY FRAMEWORK FOR THE ARAB DEAF-BLIND

The proposed system and all of its components cooperate to produce the educational material for deaf-blind in an interactive way .The proposed system is a 3-tier [50] system architecture which means Database retrieval & updating , Application Logic and GUI presentation as shown in Figure 1.

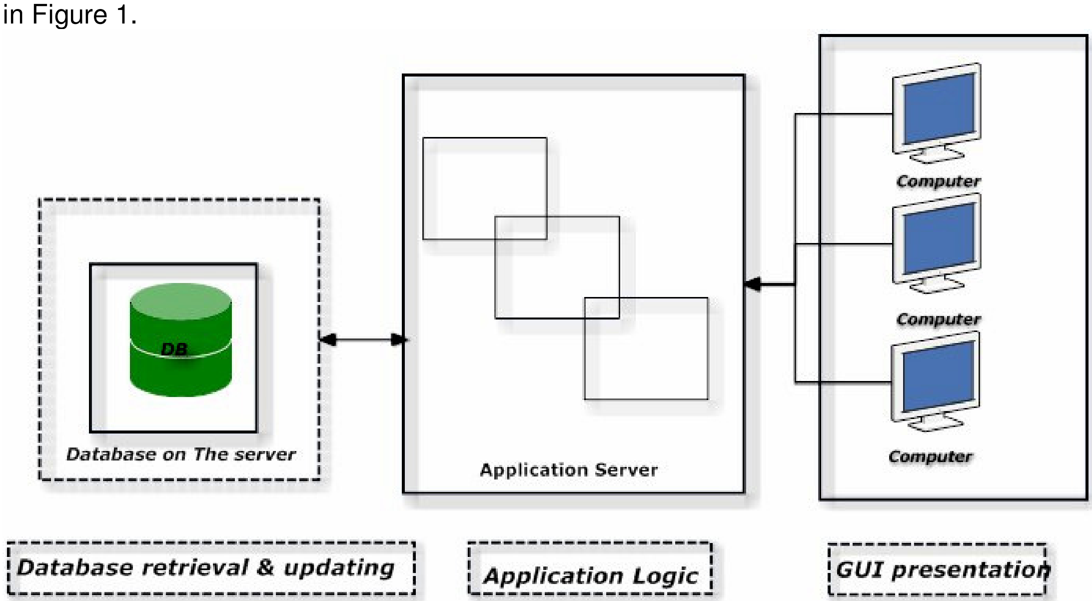



**FIGURE 1:** The 3-tier architecture of the proposed framework

Following are the proposed components of the proposed framework as shown in Figure 2.

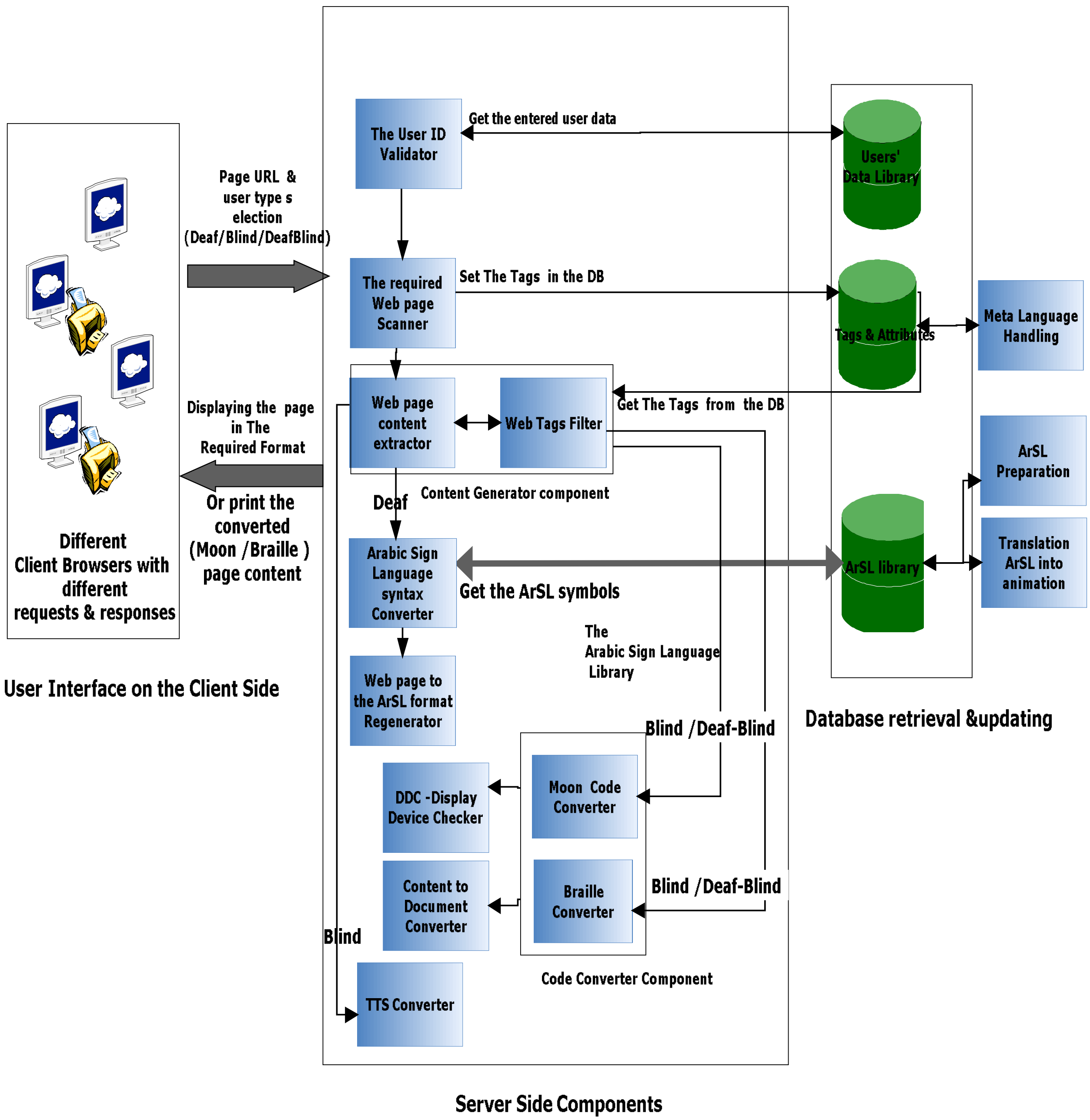



**FIGURE 2:** The block diagram architecture of the proposed framework

### 3.1 The Client Side Component

- The User Interface: Through which the user access the system first he/she enters the required URL then he/she choose the type of the disability (Deafness /Blindness /Deaf-Blindness). Then the request is processed though the server side components and the response are sent back to the user.

### 3.2 The Server Side Components

- The User ID Validator : Checks the user's login information and identify the user's disability type
- Users' Data Library :Contains all the user's information ;Name , password ,e-mail address

, disability type ,and ; if the user is blind ;blindness language (Braille /Moon ) .

•The required Web page Scanner: It's the first component in the server side of the application. It Opens the required Web page and scans the full content of it, and then the scanned content is passed to the suitable component.

•Meta Language Handling: This module is a summarization for the steps that researcher followed to be able to handle the meta-language of any Web page.

•The Tags Library : All the scanned tags from the Web page are inserted in a database to be processed later in the content generator component

•The Page Content Generator: It has two subcomponents 1-Web page content extractor & 2-Web Tags Filter.

o The Web page content extractor extracts the Arabic Web page content.

o The Web Tags Filter component eliminates each tag in the extracted content and returns with only the pure page content .Then it passes the extracted content to the suitable next components.

•Arabic SL syntax Converter : Translates the required Web page to the ArSL getting the synonymous SL words from the ArSL library

•Arabic SL Library Preparation Module: The ArSL isn't like the Arabic language [25] neither in its vocabulary nor its grammar. The first requirement to build the dictionary system was collecting the knowledge required to deal with the ArSL experts, specialists and reading different ArSl references. The second step is getting an article in a specific field and applying the ArSL knowledge on it. To get the words that can be translated into ArSL from any page, all the words must be converted to their roots.Then the whole synonymous words must be declared . Then check for ArSL synonymous existence.

The Translation of the ArSL Words into Animations Module: After collecting the words that can be represented in ArSL .The gestures of each word must be declared and stored in a database with the corresponding word. Then the animation for each word is designed from the word gestures .Finally the animations are restored in the database.

•ArSL library: This library contains of about 3500 Arabic gestures corresponding to the most common Arabic words [51] .Then the output of the ArSL syntax converter is passed to the suitable component.

•Web page to the ArSL format Regenerator: This component converts the signs to an adequate format for the Web, by regenerating the Web page content with its tags again.

•Text To Speech (TTS) converter: TTS is an external service that is used within the framework. (If the user is a blind) he/she can select whether the page is spoken loudly It converts the page content to speech. (Note: There still some problems in converting the text to some languages such as Arabic and Farsi [38, 51] or not.

• Code Converter It takes the user selected blind-language (Braille or Moon) as an input and gets the appropriate one as an output.

o Moon Code Converter: It converts the Web page content to Moon code.

o Braille Converter : :It converts the Web page content to Moon code

•Display Device Checker (DDC): If the Display Device (DD) is selected to display the output on an adequate device; the DDC checks if the suitable display device (Braille display device or Moon display device) is plugged to the computer or not .If it's plugged, the output will be transferred to it .If there is no device plugged; a special error message will appear It may be a spoken error message for the blind user and can be sensed as vibrates for the Deaf-Blind one.

•Content to Document Converter : :It takes its input from the (Moon code /Braille ) converter module and build a text file on the fly that contains the Web site content in a descriptive form.

If the user either a blind or deaf-blind, he/she can select one of the two selections –if the speaking the page loudly wasn't selected-the type of the output language, Braille or Moon code, (But there is still a problem, there is no special display device for moon characters) And select whether the output will be on a Braille /Moon display device or in a document to be printed with a special embossers later.

### 3.3 Interaction Description between the Proposed Framework Components

Flowchart shown in Figure 3 shows the interaction between the framework components.

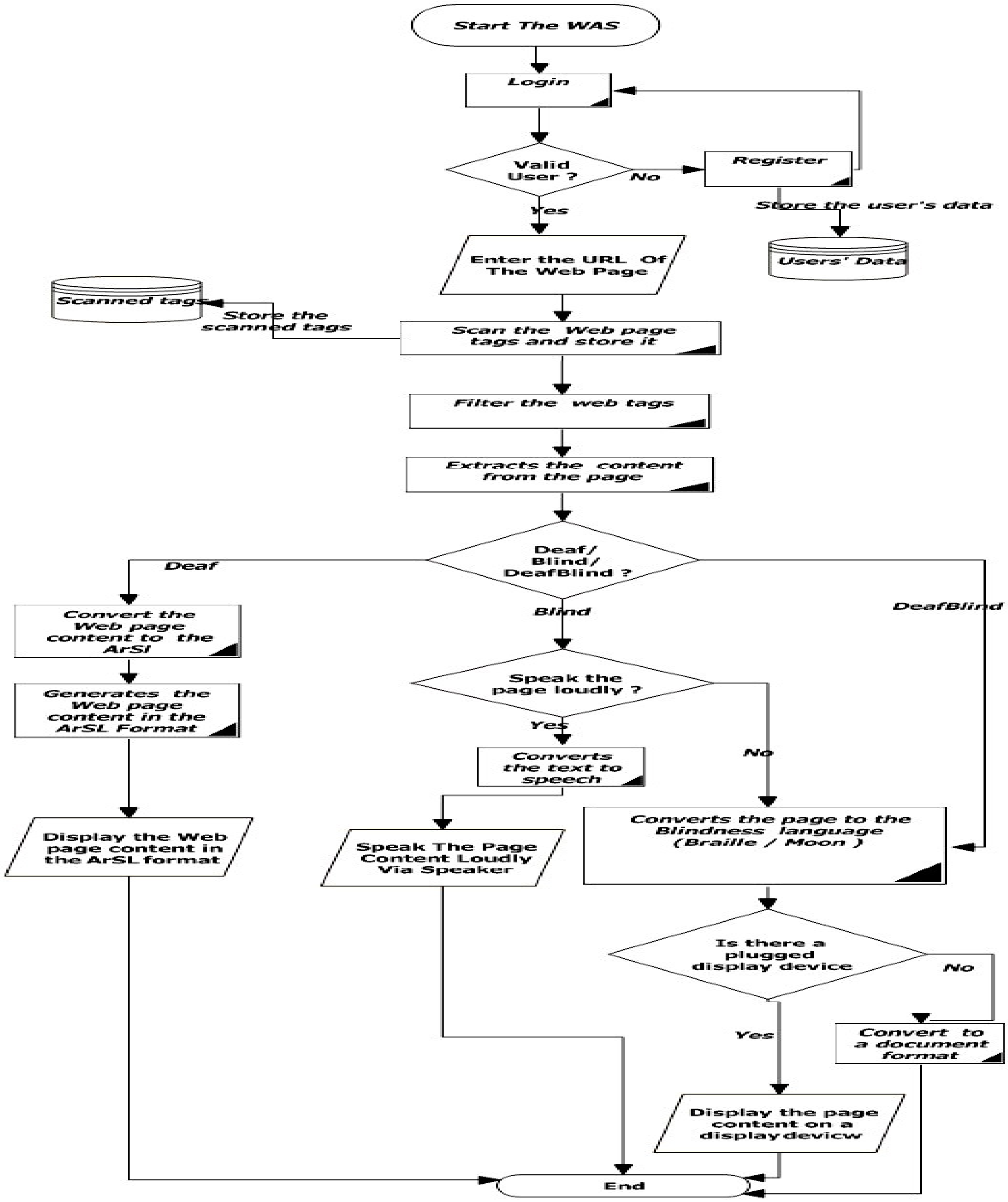



**FIGURE 3:** The system flowchart of the proposed Deaf/Blind framework

As shown in Figure.3; the user logs into the application the he/she enters the URL of the required page and selects then the page is scanned in the system to extract the page content of it .If the user is deaf, the content is converted to the SL syntax using the ArSL library The deaf user gets the page in his own language; SL. If the user is blind the content is spoken loudly using the TTS module. If the user is blind/deafblind the descriptive content is converted to either the Moon code or Braille then it's displayed on a special display device (If there is someone plugged) .If there is no device plugged, the descriptive content is converted to a document with to be printed in the client side with a special embossers.

## 4. THE PROPOSED ARABIC MOON CODE

Authors faced the problem with how can they make the Arabic websites accessible for the Arabic deaf-blind persons. They studied all the available communication methods for the deafblind people.

Depending on that study and studying the advantages of the Moon language that was mentioned in the subsection 2.2.1, authors built a new Moon code font for Arabic alphabets **TimesMoon**. They built **TimesMoon** depending on the formation of the Times new Roman font. **TimesMoon** is the first version for this new font and need many updates for the Arabic language. The reading method for this code will be from the right to the left as Arabic ,so that the Arabic deafblind user won't face problems with reading his/her mother language. Then authors used the **TimesMoon** font in the proposed framework for Arabic Deaf-Blind to facilitate the Web accessing for them. Authors used the Fontlab application to implement the design of the proposed Arabic Moon code .

Tables 1, 2 show the **TimesMoon** alphabet and its Unicode that corresponds to each Arabic letter

| Unicode | الحرف بلغه مون [illegible] | الحرف بالعربيه |
|---|---|---|
| 0627 | ı | أ ,ا |
| 0628 | ب | ب |
| 062A | ü | ت |
| 062B | ث | ث |
| 062C | ⊃· | ج |
| 062D | ⊃ | ح |
| 062E | ⊃̇ | خ |
| 062F | ـ | د |
| 0630 | ـ̇ | ذ |
| 0631 | / | ر |
| 0632 | /̇ | ز |
| 0633 | v | س |
| 0634 | v̈ | ش |
| 0635 | ک | ص |
| 0636 | ک̇ | ض |
| 0637 | ⌟ | ط |
| 0638 | ⌟̇ | ظ |
| 0639 | c | ع |
| 063A | ċ | غ |
| 0641 | ȯ | ف |
| 0642 | ō | ق |
| 0643 | ں | ك |
| 0644 | ں | ل |
| 0645 | ٩ | م |
| 0646 | ن | ن |
| 0647 | ه | ه |
| 0648 | ۶ | و |
| 0649 | ≤ | ي |
| 064A | < | ى |

**TABLE 1:** The alphabet in Arabic Moon (**TimesMoon**)

| Unicode | الارقام بلغه مون (الا/ ٩١ ب لغه ٩ و ن) | الارقام بالعربيه |
|---|---|---|
| 0023 | ÷ | (Start)# |
| 0061 | ı | ١ |
| 0062 | Г | ٢ |
| 0063 | ∨ | ٣ |
| 0064 | F | ٤ |
| 0065 | ٥ | ٥ |
| 0066 | ٦ | ٦ |
| 0067 | ٧ | ٧ |
| 0068 | ٨ | ٨ |
| 0069 | ٩ | ٩ |

**TABLE 2**: The numbers in Arabic Moon (**TimesMoon**)

## 5. CONCLUSION AND THE FUTURE WORK

This paper has been presented a We-based accessibility framework for the disabled Arab students. The proposed framework deals with three types of the disabilities: deafness, blindness and deaf-blindness. It depends on extracting the website content via reading and analyzing the meta-language of the Web page, then the extracted content is presented in a format that fits the current disabled user; printed Moon code for the deaf-blind users, ArSL animations for the deaf users or speech for the blind users.

Also a proposed e-learning system for the deaf Arab students has been presented. This system depends on the ArSL animated library. This library consists of each Arabic word and the corresponding ArSL designed animation. The animations designed for this purpose are very reliable; as they are very light animated gifs. Although the other existing BSL or ASL libraries depending on videos. So this library has many more advantages than any other existing one.

Also this paper has been presented proposed design for the Arabic Moon code for the deaf-blind Arabs.

Moreover, it presented a survey on the Web accessibility concept and definitions, goals and the current applied Web accessibility systems with the pros and cons of them.

### 5.1 Future Work

As always, there is a room for additional improvement. Further areas of research that have been identified warranting future investigation are:

- This work is intended to be included in a Quality Assurance program dedicated for the disabled Arab students. To give them their rights a the other students in good education.
- Implementing the proposed framework as a complete system and publish it online for all the Arab disabled because the proposed framework is a prototype level.
- Improving the system to read any material published on the web.
- Improving the ArSL animated library so that the deaf user can choose the preferred avatar.
- Expanding the framework to fit the other types of the significant disabilities including the cognitive disability .
- Expanding the framework to deal with AJAX-based sites and web3 sites.
- Building a complete e-learning system for the disabled persons
- Expanding the proposed dictionary for the other languages so that the deaf Arab persons can learn more than two languages .
- Designing and building the ArSL library with an icon maker for more reliability.
- Building an Arabic chat room dedicated for Arabic deaf persons.
- Building a web-based Arabic TTS library to improve the Blind module performance.
- Expanding the e-learning system for deaf Arab students with building an exam management system dedicated for them.

- Building an Arabic Moon display device to facilitate the interaction for the Arabic deaf-blinds.
- Test and validate all of the systems on their specified users.
- Including the Web common language in The proposed framework .

## 6.REFERENCES

[1] Jim Thatcher, Michael R. Burks, Christian Heilmann, Andrew Kirkpatrick, Patrick H. Lauke, Bruce Lawson, Shawn Lawton Henry, Bob Regan, Richard Rutter, Mark Urban, Cynthia Waddell ,"Web Accessibility: Web Standards and Regulatory Compliance ",Part 1 chapter1, ISBN-10: 1-59059-638-2 , ISBN-13: 978-1-59059-638-8 , July 2006

[2]Julio Miró-Borrás and Pablo Bernabeu-Soler ,"Text Entry in the E-Commerce Age: Two Proposals for the Severely Handicapped" , Journal of Theoretical and Applied Electronic Commerce Research , ISSN 0718–1876 Electronic Version ,VOL 4 ISSUE 1 ,APRIL 2009

[3] Survey of Income and Program Participation (SIPP) , Available: http://www.census.gov/sipp/ , last access : 12 /3/2010

[4] Association for Volunteer Administration , "Understanding Web accessibility", Journal of Volunteer Administration, Volume 21, pp. 36–39 , 2003 , Available: http://joeclark.org/access/webaccess/JVoluntAdmin.html , last access : 12 /3/2010

[5]S.Khwaldeh, N. Matar, Z. Hunaiti ," Interactivity in Deaf Classroom Using Centralised E-learning System in Jordan " , PGNet , ISBN: 1-9025-6016-7 ,2007 .

[6]Rosella Gennari, Ornella Mich ,"Designing and Assessing an Intelligent E-Tool for Deaf Children " , Intelligent User Interface Conference(IUI'08), Maspalomas, Gran Canaria, Spain., ACM 978-1-59593-987,January 2008 .

[7]Chee Kyun Ng ,Yeong Tat Liew , M. Iqbal Saripan ,Ahmad Fauzi Abas Nor Kamariah Noordin ,"Flexi E-Learning System: Disabled Friendly Education System", European Journal of Social Sciences – Volume 7, Number 2, 2008 .

[8]Mohammad Shirali-Shahreza , Sajad Shirali-Shahreza ,"Encouraging Persons with Hearing Problem to Learn Sign Language by Internet Websites ",8th IEEE International Conference on Advanced Learning Technologies(ICALT), 2008.

[9] PHP Website, Available: http://www.php.net , last access : 2512 /3/2010

[10]Nicoletta Adamo-Villani, John Doublestein, Zachary Martin ,"The MathSigner: an Interactive Learning Tool for American Sign Language", Proceeding of the 8th IEEE International Conference on Information Visualisation , 2004.

[11] Katja Straetz, Andreas Kaibel, Vivian Raithel, Marcus Specht , Klaudia Grote, Florian Kramer ,"An e-Learning Environment for Deaf Adults ",The ATBG-project (Aachener Testverfahren zur Berufseignung von Gehörlosen) , Aachen University ,2002.

[12] A.S.DRIGAS, "An e-Learning Management System for the Deaf people",WSEAS TRANSACTIONS on ADVANCES IN ENGINEERING EDUCATION, Issue 1, Volume 2, 2005 .

[13] James Ohene-Djan & Saduf Naqvi ,"An Adaptive WWW-based System to teach British Sign Language", Proceeding of the 5th IEEE International Conference on Advanced Learning Technologies (ICALT'05),2005 .

[14] D. Stewart,"American Sign Language (ASL)Browser", Michigan State University, May 2007, Available: http://commtechlab.msu.edu/sites/aslweb/ , last access : 12 /3/2010

[15] Apple Website , Available: http://www.apple.com/ quicktime , last access : 12 /3/2010

[16] Deafy Website , Available: http://www.deafy.net , last access : 12 /3/2010

[17] The Kuwait Sign Language , :Available: http://d2000.4mg.com , last access : 12 /3/2010

[18] Webducation Software Planungs and EntwicklungsgmbH, "E-Learning fuer Gehoerlose (ELGE)", 2009, Available: http://:www.webducation.cc , last access : 12 /3/2010

[19] Eleni Efthimiou and Stavroula-Evita Fotinea , "An Environment for Deaf Accessibility to Educational Content", Greek national project DIANOEMA ,(GSRT, M3.3, id 35), Hammamet-Tunisia , ICTA,2007 .

[20] Elizabeth Burd ,Joanna Donkin, Sarah Marshall , Cornelia Boldyreff ,"The Case for the Use of Plain English to Increase Web Accessibility", 3rd International Workshop on Web Site Evolution IEEE WSE, 2001

[21]Tirthankar Dasgupta ,Sambit Shukla,SandeepKumar;Synny Diwakar;Anupam Basu , " A Multilingual Multimedia Indian Sign Language DictionaryTool", The 6th Workshop on Asian Languae Resources, 2008

[22]Thomas Troelsgård and Jette Hedegaard Kristoffersen ,"An electronic dictionary of Danish Sign Language" ,Theoretical Issues in Sign Language Research Conference, Florianopolis, Brazil, December 2006. (2008)

[23]Chiara Vettori, Oliver Streiter, Judith Knapp ," From Computer Assisted Language Learning (CALL) to Sign Language " , Processing: the design of e-LIS, an Electronic Bilingual Dictionary of Italian Sign Language and Italian , Fourth International Conference on Language Resources and Evaluation. Post-conference workshop. Paris: ELRA. ,2004

[24]Di Mascio Tania Gennari Rosella ,"An Intelligent Visual Dictionary For Italian Sign Language ",Journal Of Web Engineering, Vol. 7, No. 4 ,318, Rinton Press ,2008

[25]Mohamed Mohandes ," Automatic Translation of Arabic Text to Arabic Sign Language ", AIML Journal, Volume (6), Issue (4), December, 2006 , Available: http://www.icgst.com/aiml/Volume6/Issue4/P1120615113.pdf , last access : 12 /3/2010

[26] Emiko Suzuki and Kyoko Kakihana ," Japanese and American Sign Language Dictionary System for Japanese and English Users " , Institute of Linguistics, Academia Sinica , 2005

[27]Martindale's Language & Translation Center Available: http://www.martindalecenter.com/Language_6_SignLanguage.html , last access : 12 /3/2010

[28] A Guide to British Sign Language, Available: http://www.britishsignlanguage.com/ , last access : 12 /3/2010

[29] German Sign Language , Available: http://www.german-language-school.info/german/signlanguage.asp , last access : 12 /3/2010

[30] British Sign Language, BSL Dictionary learn online , Available: http://www.learnbsl.org/ , last access : 12 /3/2010

[31] American Sign Language ,Available: http://www.yourdictionary.com/languages/sign.html , last access : 12 /3/2010

[32] ASL ,Available: http://www.aslpro.com/cgi-bin/aslpro/aslpro.cgi , last access : 12 /3/2010

[33] American Sign Language Dictionary, Available: http://www.handspeak.com/tour/index.php , last access : 12 /3/2010

[34] American Sign Language browser , Available: http://commtechlab.msu.edu/Sites/aslweb/browser.htm , last access : 12 /3/2010
[35] Fachgebärdenlexikon Psychologie , Available: http://www.sign-lang.uni-hamburg.de/projekte/psychlex.html , last access : 12 /3/2010

[36] Sign-Language Translator ,Available: http://www.technologyreview.com/communications/21944/?a=f , last access : 12 /3/2010 , last access : 12 /3/2010

[37] SignPuddle , Available: http://www.signbank.org/SignPuddle1.5/searchword.php?ui=1&sgn=5&sid=352&search=why+signwriting&type=any , last access : 12 /3/2010

[38]National Institute For Standards and Technology , Available: http://www.nist.gov/index.html , last access : 12 /3/2010

[39] Abdelsalam (Sumi) Helal, Mounir Mokhtari, Bessam Abdulrazak , "Engineering Handbook of Smart Technology for Aging, Disability and Independence", John Wiley & Sons, Inc , ISBN: 9780471711551 , Online ISBN: 9780470379424 , 2008.

[40]Hiroshi HORIO and Chikamune ,"Investigation of thermal characteristics of palm in order to develop a communication device for the Deaf-blind" ,WADA SICE-ICASE International Joint Conference , Bexco, Busan, Korea ,2006 .

[41]Yu-Fen Fu and Cheng-Seen ,"Static Finger Language Recognition for Handicapped Aphasiacs", Ho 0-7695-2882-1/07 IEEE ,2007 .

[42]Deafblindness, Available: http://www.tiresias.org/accessible_ict/deafblindness.htm , last access : 12 /3/2010

[43] IDEA Federal Definition of Deafblindness , 2004 ,Available: http://www.nd.gov/deafblind/ last access : 12 /3/2010

[44] Marion A. Hersh and Michael A. Johnson," Assistive Technology for the Hearing-impaired Deaf and Deafblind "; Springer London ; 28 April 2006

[45] Hearing Loss A growing problem that affects quality of life , National Academy On An Aging Society , December 1999 .

[46] William Moon Webb pages , Available: http://www.scip.org.uk/moon/homepage.htm , last access : 10 /1/2010

[47] World Health Organization , Available: http://en.wikipedia.org/wiki/World_Health_Organization , last access : 12 /3/2010

[48] The Screen Magnifiers Homeb , Available: http://www.magnifiers.org/ , last access : 12 /3/2010

[49] Comparison of screen readers - Wikipedia, the free encyclopedia , Available: http://en.wikipedia.org/wiki/Comparison_of_screen_readers , last access : 12 /3/2010

[50] Channu Kambalyal , 3-Tier Architecture , Available: http://channukambalyal.tripod.com/NTierArchitecture.pdf , last access : 12 /3/2010

[51] The Arabic Sign Language Dictionary , Available: http://www.menasy.com/ , last access : 12 /3/2010

# Computer Input with Human Eyes-Only Using Two Purkinje Images Which Works in a Real-Time Basis without Calibration

**Kohei Arai** arai@is.saga-u.ac.jp
*Department of Information Science*
*Saga University*
*1 Honjo, Saga, 840-8502, Japan*

**Makoto Yamaura** makoto@ip.is.saga-u.ac.jp
*Department of Information Science*
*Saga University*
*1 Honjo, Saga, 840-8502, Japan*

**Abstract**

A method for computer input with human eyes-only using two Purkinje images which works in a real time basis without calibration is proposed. Experimental results shows that cornea curvature can be estimated by using two light sources derived Purkinje images so that no calibration for reducing person-to-person difference of cornea curvature. It is found that the proposed system allows users' movements of 30 degrees in roll direction and 15 degrees in pitch direction utilizing detected face attitude which is derived from the face plane consisting three feature points on the face, two eyes and nose or mouth. Also it is found that the proposed system does work in a real time basis.

**Keywords:** computer input with human eyes only, cornea curvature, eyeball rotation angle estimation, Purkinje image

## 1. INTRODUCTION

Now a day, several techniques for measuring eye movement behaviour are available [1],[2],[3]. Also "double-Purkinje-image (DPI) eye tracking system is proposed [4]. This technique is based on capturing reflected infrared light that is projected on the eye. Other than these, there are Leuven dual-PC eye-tracking system [5],[6]. As for the computer input system with human eyes only based on an image-analysis method, many methods have being proposed so far. Matsuda et al. makes the line of sight which connects eyeball rotation center coordinates and a pupil center, and is performing gaze measurement [7]. Eyeball rotation center coordinates are searched for by moving an eyeball in the various directions before gaze measurement. Therefore, since there is no necessity of showing an index, gaze measurement is possible in all places, but preparation takes time and there is a fault of not permitting head movement. Moreover, Ono et al. makes the line of sight which connects cornea center-of-curvature coordinates and a pupil center, and is performing gaze measurement [8]. Ebisawa, et. al. proposed 3D eyes detection together with gaze estimation with two cameras for acquiring 3D data so that users' movement can be estimated [9]. The proposed system allows users' movement. Meanwhile, Tanaka et. al. proposed double Purkinje method for eye detection and tracking [10]. All these methods and systems need some time consumable calibration process which allows geometric relation among eyes, display and cameras.

Cornea center-of-curvature coordinates are the light source of one point, and installing a camera on the optical axis, and are searched for using the general cornea radius-of-curvature value. However, since they assumed models, such as Japanese typical eyeball form, the error remained in the direction estimation of a gaze not a little, and if these methods did not perform a calibration, when there were, they did not become. That is, in order to cancel the gaze estimation error based on the gap of a central fovea to an eyeball center, the refraction in a cornea, and the individual difference concerning the form of a cornea, the calibration which draws a gaze correction coefficient needed to be performed by gazing at two or more indices displayed on the display one by one. Moreover, since it was not what permits a motion of a user, the burden has been forced upon the user. All these systems insist users to fix their face. More than that, these systems need some time consumable calibrations in advance to use the syetms in order to make sure that the gaze locations.

The former presumes the point of regard on the display at the point of the look obtained from gaze estimation (three dimension measurement) of not only the direction of a gaze but both eyes, and uses the pupil center and the corneal reflex center for gaze estimation. The latter presumes a cornea center of curvature using two Purkinje images, makes a user gaze at the index of three points, presumes an eyeball rotation center, is the method of making a gaze the straight line which connects these, and is verifying accuracy by the experiment using a model eye. In this paper, an expensive stereo camera is not needed, but only a cheap simple eye camera permits a motion of a user, and the method of determining the direction of a look from a pupil center and a cornea center of curvature is proposed without the calibration which forces a user a gaze of three points. By specifically measuring an eyeball cornea curvature radius simply, the degree estimation of eyeball rotation angle which does not need a calibration is performed, details are extracted from a face picture, the posture of a head is detected from those relative spatial relationships, and a motion of a head is permitted. The light source of two points was used for measurement of the cornea curvature radius of an eyeball, and two Purkinje images obtained from the cornea surface were used for it. At this time, it decided to also use together the near-infrared light source which a camera has using the near-infrared camera which became budget prices, and to acquire the clear Purkinje image in recent years. When five subjects estimate the direction estimation accuracy of a look for this, a motion of the head of the 30 roll directions and the 15 directions of a pitch is permitted, and since it checked that the direction of a gaze could be presumed without a calibration with the error of 0.57 to 0.98 degrees, it reports here.
The angle estimation method of eyeball rotation angle of using two light sources and one near-infrared camera is proposed first, the head angle detection method using the details in a face picture is described, and this paper estimates the validity of the proposed method by the direction estimation experiment of a look by five subjective examiners.

## 2. PROPOSED METHOD

Angle estimation of eyeball rotation angle and direction estimation method of look is as follows, originally, a look is a straight line which connects a central fovea, the node of a lens, and the point of regard, as shown in Figure 1, but since these position specification is difficult, it considers the straight line going through a cornea center of curvature and a pupil center to be a look. Four kinds of Purkinje images used for look estimation exist, as shown in Figure 2, but in this paper, the 1st two Purkinje images with the biggest amount of catoptric light from two light sources are used. It projects on right and left in every one front of the Purkinje cornea from a camera optic-axis center. The example is shown in Figure 3. Next, the procedure of presuming the point of regard on a display from two Purkinje images is shown.

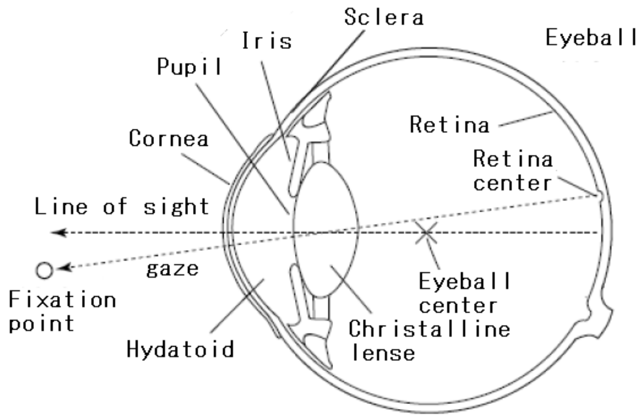



**FIGURE 1:** Definition of eye gaze and line of sight

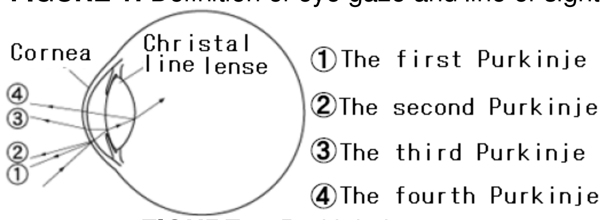



**FIGURE 2:** Purkinje images

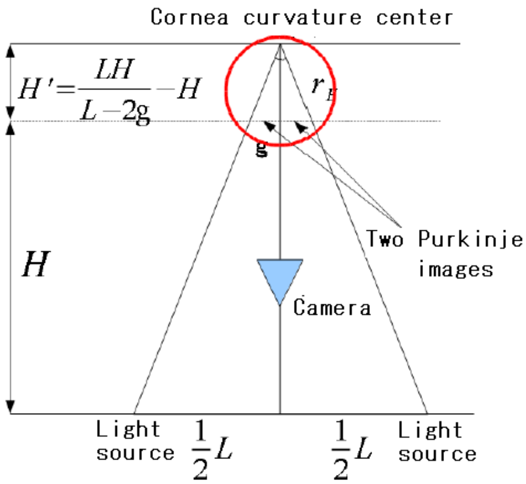
**FIGURE 3:** An example of Purkinje image

First, as shown in Figure 4 (a), the geometric spatial relationships of the Purkinje image on two a light source, one camera, and a simple eye cornea is considered. Two Purkinje images exist on the line which divides into two equally the angle which the straight line which connects a cornea center of curvature to a light source and a camera, respectively makes. Extraction of the Purkinje image and the distance between two Purkinje images are calculable with image processing. Since the Purkinje image exists as catoptric light on a cornea, compared with the circumference, a luminosity value becomes remarkably high. Therefore, if there are some which show a high luminosity value, let this be the candidate of the Purkinje image. Moreover, since near a pupil exists, the Purkinje image extracts two candidates Purkinje who exist in the position nearest to a pupil. Purkinje is made into the distance between Purkinje, the center of gravity of two obtained candidates. This process is shown in this Figure (c).
In Figure 4 (a), a formula (1) can express a cornea curvature radius $r_E$.

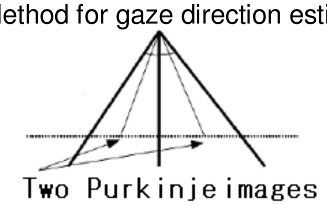
$$r_E=\sqrt{\left(\frac{1}{2}\cdot 2g\right)^2+\left(\frac{lH}{L-2g}-H\right)^2} \qquad (1)$$

where L denotes the distance between two light sources, H means the distance between camera and eyeball surface, and g denotes the measured distance between two Purkinje images. Thus, a cornea curvature radius is searched for.
Next, a pupillary zone is extracted using a dark pupil method, ellipse approximation of the pupil form is carried out, and a pupil center is searched for. With a dark pupil method, if an infrared floodlight is installed in the position distant from the lens optic axis of a camera and an eyeball is illuminated, the reflection of light of the pupil portion of the illuminated eyeball will be lost, and the pupil portion of the eyeball picture caught with the camera will use the character to become dark here. Image processing extracts this dark portion. Moreover, in the extracted pupillary zone, since

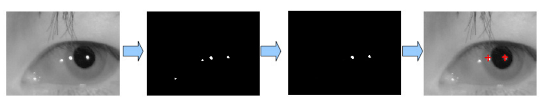
reflected lighting and the up-and-down portion of a pupil hide and are missing by the eyelid, the pupil may not be able to be extracted correctly. Therefore, edge is detected from the obtained temporary pupillary zone, and the edge concerned is approximated to an ellipse by a least-squares method. However, when it calculates using except a pupil outline, there is a possibility of resembling the mistaken ellipse. Therefore, pupil ellipse approximation using the character of an ellipse was performed. That is, in Figure 5, straight lines l, m, and n shall be parallel, and l and n shall exist in the equal distance from m.




(a) Method for gaze direction estimation




(b) Estimation of cornea curvature radius (Purkinje image is situated on the bisector between light source and cornea curvature center



(c) Process flow of the method for gravity center estimation using two candidates of the detected Purkinje images

**FIGURE 4:** Method for estimation of cornea curvature radius using the distance between two Purkinje images.

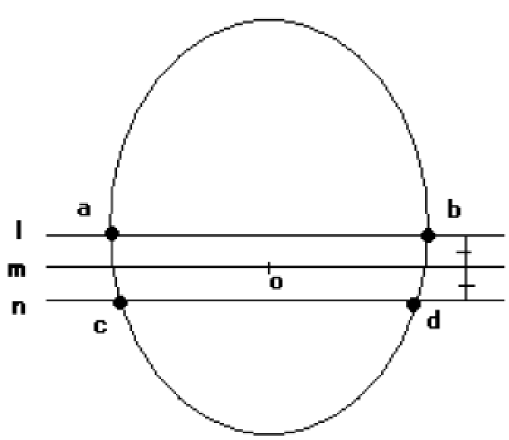



**FIGURE 5:** Extraction of sample points from the assumed ellipsoid.

The intersection of an ellipse and a straight line l is set to a and b, the intersection of an ellipse and a straight line n is set to c and d, and the middle point of the intersection of an ellipse and a straight line m is set to o. When the middle point of a line which connects the middle point of a and b and the middle point of c and d is set to $o_0$, there is character in which $o_0$ overlaps with o.
N which is in the equal distance about a temporary pupillary zone from the straight line m drawn in the center. The point of hitting $o_{i0}$ (i = 1-N) of the group of the parallel lines is searched for. The point searched for is distributed on a straight line m. When there are few noises enough, the position in which most many points o0 gathered is equivalent to the position of o. Since the point which is distant from the position will include the point which is not on the locus of an ellipse, it accepts them. o point $a_i$ which is alike and has sufficiently near $o_{i0}$; $b_i$; $c_i$; Since $d_i$ is a point on an ellipse, ellipse approximation of an exact pupil outline can be performed by using them. In addition, the Canny filter [11] was used for edge detection. A pupil outline can be extracted by approximating an ellipse from the pupil ellipse sample point acquired by the above-mentioned method. The process of old processing is shown in Figure 6.

(a) Original image (b) Binalized image (c) Noise reduced image (d) Edge detected (e) Extracted samples (f) Approximated ellipsoid

**FIGURE 6:** Method for estimation of approximated ellipsoid of cornea shape.

Next, a cornea center of curvature is searched for from the spatial relationship of the acquired cornea curvature radius and the light source of Figure 7, an eyeball, and a camera, and the geometric expression of relations in a Figure. A cornea center of curvature exists here on the bisector which ties a camera, the Purkinje image, and lighting. It asks for a look vector from a pupil center and a cornea center of curvature.
Finally the point of regard on a display is computed from a look vector. The vector which passes along two points, a cornea center of curvature and a pupil center, is made into a look vector, and the point of regard on a display is computed.

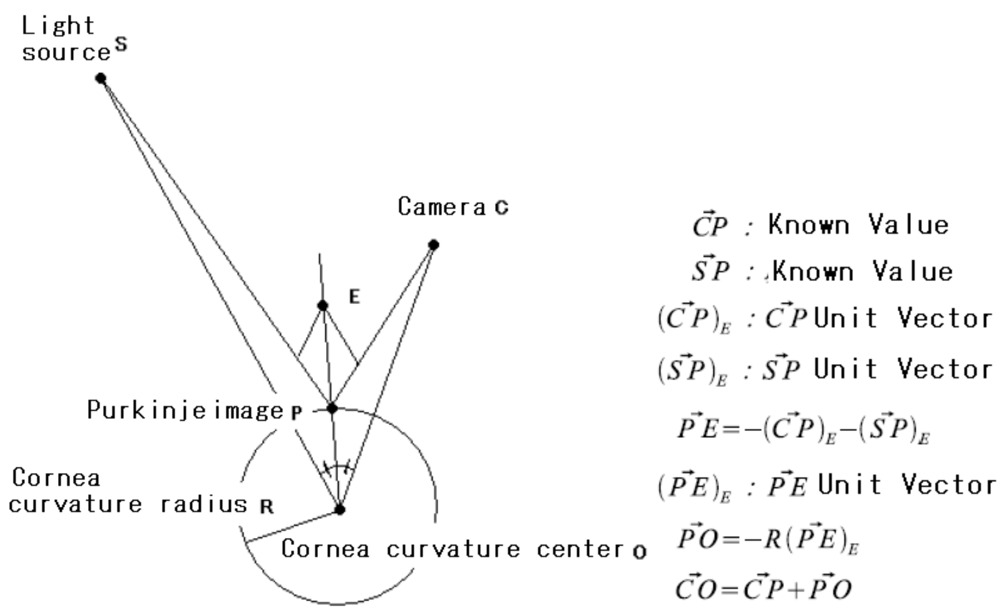


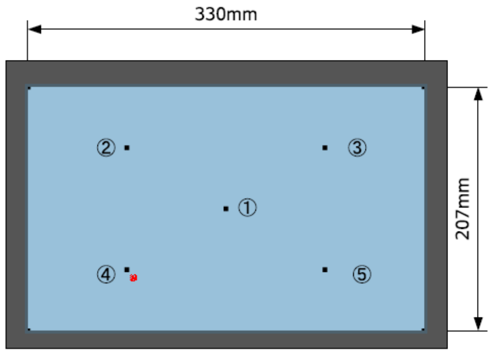

**FIGURE 7:** Geometric relation among light source, camera and eyeball

It is considered as a camera position *(0, 0, 0)*, and is considered as the main coordinates $(0, 0_z)$ of the picture picturized with the camera. If a look vector is made into $v = (x_v; y_v; z_v)$, a camera, lighting, and a display assume that it is being fixed and the head is also being fixed, and are *z*.

The distance of the direction of an axis is known. When distance of an eyeball and a display is set to $z_h$, point-of-regard coordinates $t$ = on a display $(x_t; y_t)$ is,

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} x_v \times \frac{z_h}{z_v} \\ y_v \times \frac{z_h}{z_v} \end{pmatrix} \qquad (2)$$

so that the gaze fixed point can be calculated as is shown in Figure 8.




(a) Calibration point locations on the computer screen for evaluation of gaze position estimation accuracy (red point shows the estimated gaze fixed point)

(b) The estimated cornea center and pupil center

**FIGURE 8:** Method for gaze fixed position estimation.

How to permit a motion of a head finally is shown. The both ends of eye and a mouth, the middle point between two ends of eye, and the middle point of the both ends of a mouth are detected from a face picture. Template matching is used for pursuit of details at these details extraction using OpenCV [10]. By the middle point of both eyes, and the middle point of a mouth, three square shapes can be constituted and a plane can be defined. The normal direction of this plane is judged to be the posture of a head. Once extracting these details, by using template matching which can pursue most in a short time, it is devising so that the pursuit in real time may be possible. An example of details extraction and pursuit is shown in Figure 9. The distance between the details in a Figure is found and the degree of rotation angle of a head can be presumed in comparison with the position of the details extracted in early stages by OpenCV. The blue line segment of Figure 9 is a head posture (the plane normal direction is shown.). Moreover, a green line segment connects the middle point between the middle points of both eyes, and the middle point of the both ends of a mouth.

The geometric relation between the rotation center of a head and an eyeball rotation center is shown in Figure 10, and can presume an eyeball rotation center position after this. If the method of presuming a look from an eyeball rotation center position is considered to be the same thing as the above-mentioned method, a motion of a head will be permitted and the direction estimation of a look will be attained. At this time, that the both ends of both eyes and the both ends of a mouth are not occluded are the conditions which presume a head rotation center and an eyeball rotation center, and 30 degrees and about 15 degrees are rotation allowable limits in the roll direction and the direction of a pitch.

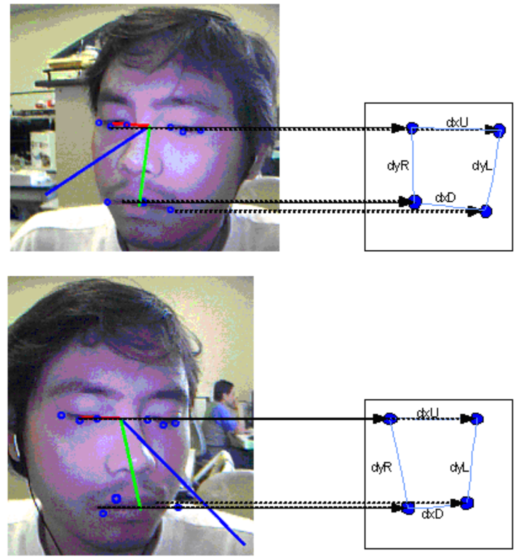



**FIGURE 9:** Method for head pose angle estimation using the extracted feature points from the acquired face image.

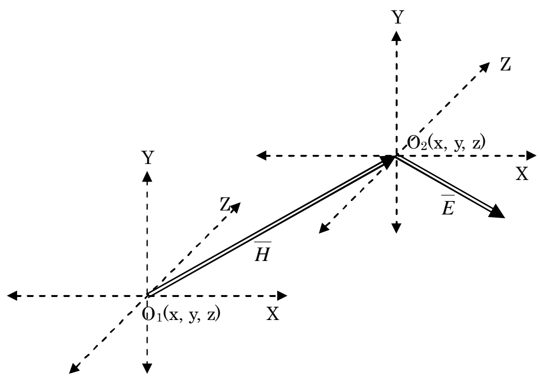



**FIGURE 10:** Geometric relation between two coordinates of head and eyeball

## 3. EXPERIMENTS

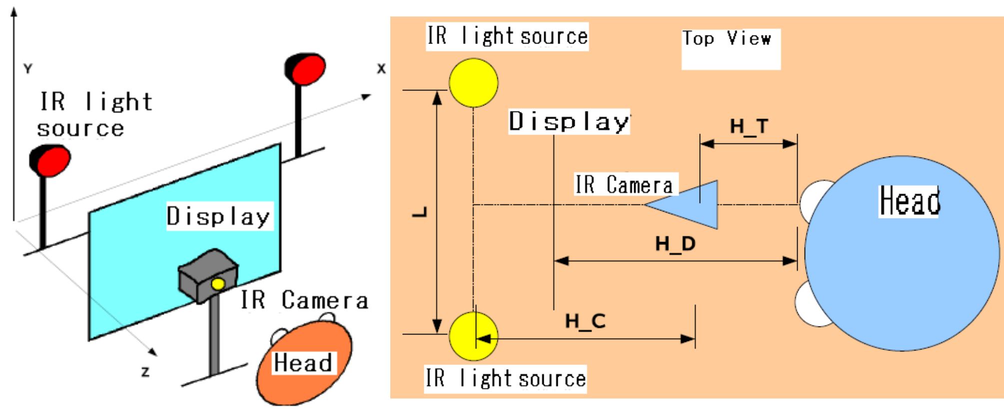
### 3.1 System Configuration

The measurement equipments used is the followings,
PC: Dell Computer Optiplex 755 Core 2 Quad 2.66 GHz CPU with 2 MB RAM of main memory
OS: WindowsXP home Service Pack2IR
Camera: (640 by 480 pixels) of DC-NCR 131 type manufactured by NetCowBoy and frame rate: 10 frames-per-second
Infrared floodlight: KMT-7787.
These measurement equipments are installed as shown in Figure 11.

### 3.2 System Parameters

The parameters of the experimental conFigureuration are as follows,
$L = 1000$ mm; $H_C = 670$ mm; $H_T = 150$ mm; $H_D = 380$mm
Software development environment is as follows,
Microsoft Visual C++

Microsoft Visual studio.NET2
OpenCV 1.0
Moreover, the picture acquired from the camera on the occasion of look measurement was processed in real time, and performed accuracy verification.




(a) Experimental configuration (b) Experimental layout from the top view
**FIGURE 11:** Experimental layout and configuration

### 3.3 Cornea Curvature Measurements

The measurement result of a cornea curvature radius is shown in Figure 12.

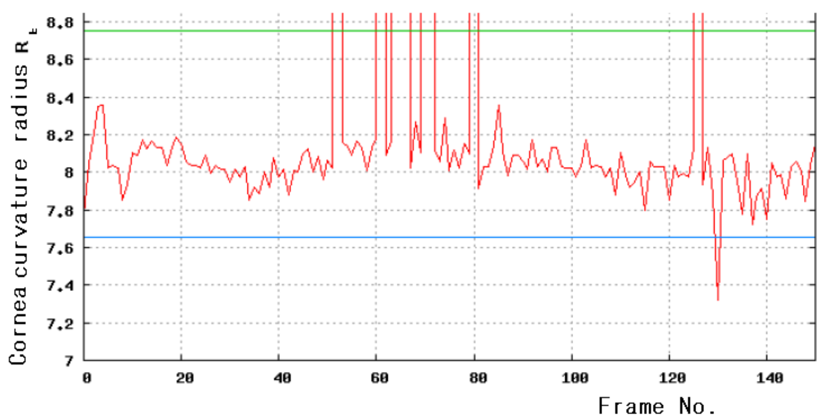



**FIGURE12:** An example of experimental results of cornea curvature radius estimation.

The spike in a Figure is based on incorrect detection of the Purkinje image, and these can be accepted from measurement of a cornea radius. Moreover, two lines in a Figure are the maximum errors of the cornea curvature-radius point estimate considered when movement of ± 1.0cm has a head in the display direction. The average of the cornea curvature-radius point estimate except an edge portion is used for the point estimate of the cornea curvature radius used for the actual degree estimation of eyeball rotation angle. The value of a cornea curvature radius was set to $R$ = 7.92 mm from this result.

### 3.4 Eyeball Rotation Angle Estimation Accuracy

In the experiment, eyeball angle estimation accuracy was checked by seeing five indices displayed on the display one by one, where a head is fixed simply. The presumed result of the eyeball angle detection from which a cornea curvature radius differs is shown in Table 1. It is delta $x$ and $y$ when a cornea curvature radius is set to $R$ = 7.92 mm at this time. $\Delta xy$ = it is 1.128.

| Radius | Mean Δx | Mean Δy | Mean Δxy |
|---|---|---|---|
| 7.5mm | 0.7762 | 1.4836 | 1.6743 |
| 7.6mm | 0.5432 | 1.1397 | 1.2626 |
| 7.7mm | 0.4246 | 1.0819 | 1.1622 |
| 7.8mm | 0.4347 | 1.0324 | 1.1202 |
| 7.9mm | 0.5283 | 0.9889 | 1.1211 |
| 8mm | 0.6739 | 0.9504 | 1.1651 |
| 8.1mm | 0.8462 | 0.9166 | 1.2475 |

**TABLE 1:** Mean of error in eyeball rotation angle estimation for the various cornea curvature radius (five people of examiners)

Near the measured cornea curvature radius has the highest degree detection accuracy of eyeball rotation angle. Therefore, it has checked that rather than led which uses the measured cornea curvature radius to improvement in the degree detection accuracy of eyeball rotation angle using the standard value of a cornea curvature radius.
Change of the look estimation result of survey by head parallel translation and rotation is shown in Table 2.

(a) Head translation in x direction

| Traslation in x | Mean Δx | Mean Δy | Mean Δxy |
|---|---|---|---|
| -2.0cm | 0.8123 | 1.00 | 1.29 |
| -1.0cm | 0.6677 | 0.99 | 1.19 |
| 0cm | 0.5283 | 0.99 | 1.12 |
| 1.0cm | 0.6266 | 0.97 | 1.15 |
| 2.0cm | 0.7238 | 0.97 | 1.21 |

(b) Head translation in y direction

| Traslation in y | Mean Δx | Mean Δy | Mean Δxy |
|---|---|---|---|
| -2.0cm | 0.5611 | 1.65 | 1.74 |
| -1.0cm | 0.5628 | 1.27 | 1.39 |
| 0cm | 0.5283 | 0.99 | 1.12 |
| 1.0cm | 0.5620 | 1.02 | 1.17 |
| 2.0cm | 0.5500 | 1.07 | 1.21 |

(c) Head translation in z direction

| Traslation in z | Mean Δx | Mean Δy | Mean Δxy |
|---|---|---|---|
| 150mm | 0.5283 | 0.99 | 1.12 |
| 200mm | 6.0704 | 2.09 | 6.42 |
| 250mm | 10.7227 | 3.00 | 11.13 |
| 300mm | 13.8139 | 3.51 | 14.25 |

(d) Head rotation in roll direction

| Rotation Angle | Mean Δx | Mean Δy | Mean Δxy |
|---|---|---|---|
| -45deg | 1.1904 | 1.95 | 2.29 |
| -30deg | 0.9993 | 1.67 | 1.95 |
| 0deg | 0.5283 | 0.99 | 1.12 |
| 30deg | 1.2271 | 1.12 | 1.66 |
| 45deg | 1.5640 | 1.2 | 1.97 |

(e) Head rotation in yaw direction

| Rotation Angle | Mean Δx | Mean Δy | Mean Δxy |
|---|---|---|---|
| -45deg | 1.1904 | 1.95 | 2.29 |
| -30deg | 0.9993 | 1.67 | 1.95 |
| 0deg | 0.5283 | 0.99 | 1.12 |
| 30deg | 1.2271 | 1.12 | 1.66 |
| 45deg | 1.5640 | 1.2 | 1.97 |

**TABLE 2:** Eyeball rotation angle estimation errors (Mean of estimation errors for five examiners)

Here, movement and rotation (circumference of y and a z axis) for all directions of x, y, and z were investigated, respectively. When Table 2 (a) is seen, it turns out that the error generated by head movement of a x direction affects only a x direction. It is the same thing as is shown in Table 2 (b) and (c), it turns out that the error generated by head movement of y and the direction of z affects it only in y and the direction of z. This cause is because depth is given as a constant, and in order to solve this, the distance between a head and a camera just measures it on real time. If Table 2 (d) is seen, the fall for presumed accuracy with big head rotation to the roll direction will be seen. Moreover, in Table 2 (e), the big fall for presumed accuracy was expected for the head rotation to the direction of yaw not to have the influence of head rotation looked at by presumed accuracy, but to exceed a certain range in a certain fixed range. This is considered to originate in the cornea form of the human being that a cornea curvature radius becomes large steeply by the nose side gently-sloping by the ear side.

In measurement of a cornea curvature radius, it seems that the big accuracy fall has occurred when this is exceeded since it has measured in 3.0-5.0mm from the cornea center-of-curvature part. Moreover, in the influence on the presumed accuracy by the head rotation to the direction of a pitch, since the influence of eyelashes becomes large in downward rotation of a head, pupil detection becomes difficult (measurement is impossible), and since the influence of upward rotation of a head of eyelashes decreases, its pupil detection accuracy improves. From this, the camera position at the time of measurement is understood that it is desirable to install by arrangement which looks up at a head.

### 3.5 Head Pose Angle Estimation

As 3.4 described, movement in the direction of z is compared with movement to x and a y direction, and has serious influence on eyeball angle detection. Therefore, I need to get a user to maintain distance with a computer display to some extent (separating beyond the defined distance and twisting like). As Chapter 2 described, to rotation of x of a head, the parallel translation of a y direction, a roll, a pitch, and the direction of Yaw, it is detectable. As for parallel translation, it is highly accurate and can presume correctly about 100%. Therefore, the degree of setting head rotation angle and average estimation error by five subjects were checked by experiment here. An experimental result is shown in Figure. 13. Since the Yaw rotation of a head is rotation in a head plane, a presumed error is comparatively small. The presumed error of the direction of a pitch was the largest, and it turned out that the presumed error of about 5 times is produced also in rotation of about 1.0 degrees in the pitch rotation to a direction which especially bows. Even if it is relatively stout and rotates about 30 degrees to the pitch rotation to a direction which raises the head, the presumed error is settled in about 1.7 degrees. Moreover, the presumed error over rotation of the head of the roll direction was located in that middle, and when it was ±30 roll rotations, it has checked fitting in the presumed error of less than 1.7 degrees. Furthermore, one of eyes occlusion roll rotation exceeding 40 degrees, and a presumed error becomes large rapidly. Therefore, ±30-10~+30 degrees, even if the head rotated ±30 degrees in a roll, a pitch, and each direction of Yaw, it checked fitting in the presumed error of less than 1.7 degrees to them.

### 3.6 Overall Performance

A cornea curvature radius is presumed using the Purkinje image of two points obtained by using the two-point light source proposed in this paper, the proposed method of eyeball rotation angle estimation which does not need the calibration for individual difference dissolution of cornea curvature are 0.57 degree and y to the direction of x in the state where the head was fixed in simple. It turned out that the degree of eyeball rotation angle can be presumed with the error of about 0.98 degree in a direction. Moreover, since the degree estimation method of head rotation angle using the details of the face proposed in this paper is the presumed error of about 1.7 degrees under restriction of the degree of rotation angle, it assumes that rotation of these heads and rotation of an eyeball are independent, and it is RSS: Root Sum Square is taken, it turns out that the eyeball angle estimation accuracy in the case of permitting rotation of a head is 1.988 degrees in a x direction at 1.821 degrees and a y direction.

If a look is presumed and rotation is permitted for a head in the above-mentioned restriction range from the face picture acquired from the camera which separated 150mm, the direction of a look can be presumed in the error of about 2 times. When an intersection with a look is put on the computer display which left this 300mm, it is equivalent to about 10.472mm look and a display intersection position estimation error. When a with a pixel size (4 pixels / 10mm) computer display is assumed, this intersection position estimation error will be equivalent to about 4 pixels. Figure 14 sets up a target on 7 pixels from a center as an example of look stability, gets five subjects to gaze at the target concerned for 30 seconds, and shows what plotted the presumed intersection position. Therefore, by making distance between adjacent keys into 10 pixels or more showed that rotation of a head was permitted and the degree estimation method of eyeball rotation angle which does not need a calibration could be realized.

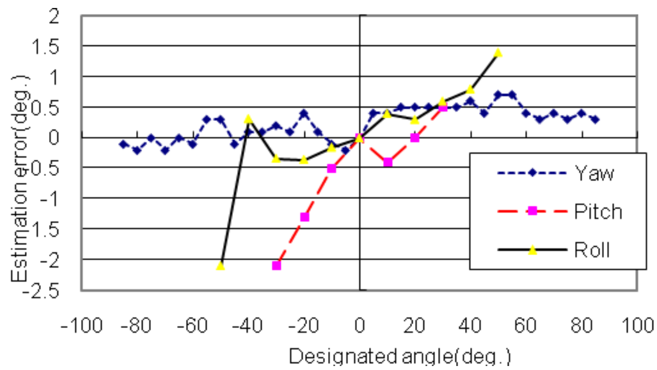



**FIGURE13:** Head rotation angle estimation error in pitch, roll and yaw directions.

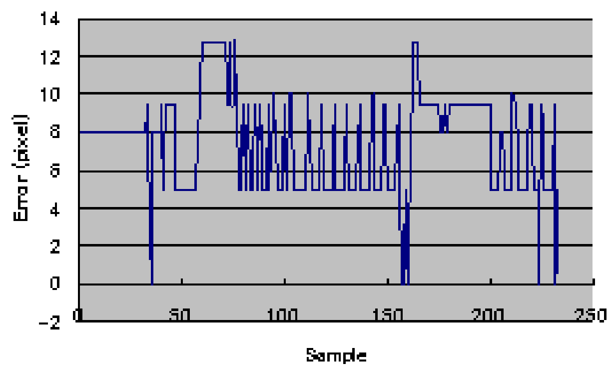



**FIGURE14:** Stability of the estimated gaze intersection on computer display. (Gaze position is set at 7 pixels from the center of the computer display as an example. The stability is within a couple of pixels mostly, -7 to +6 pixels error is happened occasionally though.)

## 4. CONCLUDING REMARKS

This paper proposed the estimation method of eyeball rotation angle which does not need the calibration for individual difference dissolution of cornea curvature by presuming a cornea curvature radius using the Purkinje image of two points obtained by using a two-point light source. As a result, in the state where the head was fixed in simple, they are 0.98deg(s) to 0.57deg(s) and the direction of y in the direction of x. The degree estimation accuracy of eyeball rotation angle was acquired with the error of a grade. Moreover, by the angle estimation method of the plane defined by the details of the face proposed since rotation of a head was permitted, when preparing restriction in each rotation of a roll, a pitch, and a yaw directions, it checked that the degree of rotation angle could be presumed with the error of about 0.5 degree. When it assumed that rotation of these heads and rotation of an eyeball were independent, it turned out that the eyeball angle estimation accuracy in the case of permitting rotation of a head is 1.988 degrees in a x direction at 1.821 degrees and a y direction. That is, since the degree of viewing angle beyond this error was detectable, when detaching and setting up the interval of a key rather than this, it turned out that it can guarantee that discernment of a key is possible.

## 5. REFERENCES

1. Carpenter, R. H. S. (1988). *Movements of the eyes*. London: Pion.
2. Collewijn, H. (1999). Eye movement recording. In R. H. S. Carpenter & J. G. Robson (Eds.), *Vision research: A practical guide to laboratory methods* (pp. 245-285). Oxford: Oxford University Press.
3. Crane, H. D., & Steele, C. M. (1985). Generation-V dual-Purkinje-image eyetracker. *Applied Optics, 24,* 527-537.
4. Deubel, H., & Bridgeman, B. (1995a). Fourth Purkinje image signals reveal eye-lens deviations and retinal image distortions during saccades. *Vision Research, 35,* 529-538.
5. Deubel, H., & Bridgeman, B. (1995b). Perceptual consequences of ocular lens overshoot during saccadic eye movements. *Vision Research, 35*, 2897-2902.
6. Van Rensbergen, J., De Troy, A., Cavegn, D., De Graef, P., van Diepen, P. M. J., & Fias, W. (1993). *The consequences of eye-lens movement during saccades for a stable retinal image.* Poster presented at the Seventh European Conference on Eye Movements. Durham, UK.
7. K. Matsuda, T. Nagami and S. Yamane, *Development of gaze measurement system,* Technical Report of the Institute of Electronics, Information and Communication Engineers, TL2002-2, 2000.
8. T. Ohno, N. Takegawa, and A. Yoshikawa, *Gaze estimation method based on eyeball shape model,* Proc. of the 8$^{th}$ Image Sensing Symposium, pp.307-312, 2002.
9. Y. Ebisawa, 3D gaze measurement equipment, Japanese Patent No.2005-198743, 2005.
10. H. Tanaka, S. Hikita, K. Kasai, and A. Takeda, *Gaze estimation method based on cornea curvature center estimation with two light sources,* Journal of the Institute of Electronics, Information and Communication Engineers, 108, 479, MBE2008-128, 117-180, 2009.
11. J. Canny: *A Computational Approach to Edge Detection,* IEEE Trans. Pattern Analysis and Machine Intelligence, .8, 6, .679-698, 1986.
12. Intel Crop.: OpenCV (Inte Open Computer Vision Library), http://www.intel.com/technology/computing/opencv/

# CALL FOR PAPERS

**Journal**: International Human Computer Interaction (IJHCI)
**Volume:** 2 **Issue:** 1
**ISSN:**2180-1347
**URL:** http://www.cscjournals.org/csc/description.php?JCode=IJHCI

## About IJHCI

The International Journal of Human Computer Interaction (IJHCI) publishes original research over the whole spectrum of work relevant to the theory and practice of innovative and interactive systems. The journal is inherently interdisciplinary, covering research in computing, artificial intelligence, psychology, linguistics, communication, design, engineering, and social organization, which is relevant to the design, analysis, evaluation and application of human computer interactive systems.

IJHCI addresses the cognitive, social, health, and ergonomic and other aspects of interactive computing including cognitive ergonomics, collaborative design, ergonomics & human factors, legal, ethical & social aspects of IT, human computer interaction, ergonomics, human computer interaction and user interface. IJHCI emphasizes the human element in relation to the systems in which humans function, operate, network, and communicate, including software, hardware, and their various contexts of use.

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJHCI.

## IJHCI List of Topics

The realm of International Journal of Human Computer Interaction (JHCI) extends, but not limited, to the following:

- Agent models Co-ordination and communication
- Agent-based Computing
- Computer Mediated Communication
- Computer Supported Cooperative Work
- Design and Evaluation of Innovative Interactive Sy
- Empirical Studies of User Behaviour
- Graphic Interaction
- Human language Technologies and Machine Learning i
- Human-Computer Interaction Theory
- Innovative Interaction Techniques

- Intelligent Tutoring, Coaching and Debugging Syste
- Interaction in Mobile and Embedded Systems
- Interface Design and Evaluation Methodologies
- Mixed and Augmented Reality
- Natural language interaction
- Presence
- User Interface Prototyping for Interactive Systems
- Intelligent user Interfaces
- Interactive Decision Support Systems
- Knowledge Acquisition, Discovery, Modeling and Man
- Multimodal Interaction
- Peer to Peer Communication Between Intelligent Sys
- Speech Interaction
- Virtual Reality

## Important Dates

**Volume:** 2
**Issue:** 1
**Paper Submission:** January 31, 2011

**Author Notification:** March 01, 2011

**Issue Publication:** March / April 2011

# CALL FOR EDITORS/REVIEWERS

CSC Journals is in process of appointing Editorial Board Members for ***International Journal of Human Computer Interaction (IJHCI).*** CSC Journals would like to invite interested candidates to join **IJHCI** network of professionals/researchers for the positions of Editor-in-Chief, Associate Editor-in-Chief, Editorial Board Members and Reviewers.

The invitation encourages interested professionals to contribute into CSC research network by joining as a part of editorial board members and reviewers for scientific peer-reviewed journals. All journals use an online, electronic submission process. The Editor is responsible for the timely and substantive output of the journal, including the solicitation of manuscripts, supervision of the peer review process and the final selection of articles for publication. Responsibilities also include implementing the journal's editorial policies, maintaining high professional standards for published content, ensuring the integrity of the journal, guiding manuscripts through the review process, overseeing revisions, and planning special issues along with the editorial team.

A complete list of journals can be found at http://www.cscjournals.org/csc/byjournal.php. Interested candidates may apply for the following positions through http://www.cscjournals.org/csc/login.php.

*Please remember that it is through the effort of volunteers such as yourself that CSC Journals continues to grow and flourish. Your help with reviewing the issues written by prospective authors would be very much appreciated.*

Feel free to contact us at coordinator@cscjournals.org if you have any queries.

# Contact Information

**Computer Science Journals Sdn BhD**
M-3-19, Plaza Damas Sri Hartamas
50480, Kuala Lumpur MALAYSIA

Phone: +603 6207 1607
+603 2782 6991
Fax: +603 6207 1697

**BRANCH OFFICE 1**
Suite 5.04 Level 5, 365 Little Collins Street,
MELBOURNE 3000, Victoria, AUSTRALIA

Fax: +613 8677 1132

**BRANCH OFFICE 2**
Office no. 8, Saad Arcad, DHA Main Bulevard
Lahore, PAKISTAN

**EMAIL SUPPORT**
Head CSC Press: coordinator@cscjournals.org
CSC Press: cscpress@cscjournals.org
Info: info@cscjournals.org

COMPUTER SCIENCE JOURNALS SDN BHD
M-3-19, PLAZA DAMAS
SRI HARTAMAS
50480, KUALA LUMPUR
MALAYSIA