Proceedings of the 10th WSEAS International Conference on COMPUTERS, Vouliagmeni, Athens, Greece, July 13-15, 2006 (pp1147-1152)

A Creation of Voice Tour Guide System by Cell Phones

Nariaki Kato
Department of Business Administration
Aichi Sangyo University
Okacho, Okazaki, Japan 444-0005

Naohiro Ishii
Department of Information Science
Aichi Institute of Technology
Yakusacho, Toyota, Japan 470-0392

Abstract: - Nearly 70% of the Japanese population has a cell phone. Cell phones are not just used to call. They are becoming a tool to search information about culture and entertainment through the Internet. At the same time, walking is popular with a health boom.

We created a voice guide system for historic sites on trial. From a statistical data, we studied trends of cell phone users and the cellular technique in the future. Also, we picked out the technical and management problems of the trial system from a questionnaire and studied the effectiveness of the system. In this paper, we describe this system would be useful if the economic problem of downloading data is solved.

Key-Words: - Cell phones, Voice, Music player, Downloading time, Economic efficiency.

1 Introduction

1.1 Background of Research

According to *White Paper on Information and Communications, 2005* issued by the Ministry of Internal Affairs and Communications, nearly 70% of the Japanese population has a cell phone. In addition, the cell phones have high-performance such as a music player and Internet connection. People hope to use the high-functions on and on[1].

At the same time, walking and hiking are popular with a health boom. The Japan Walking Association says the population of walkers in 2005 is over 40 million. Okazaki city in Aichi prefecture where I live is no exception, and many walkers gather in historic sites during the holidays. They read the guidebooks and the statements of the historic sites eagerly. That is they try to have both good health and culture.

The explanatory materials of the historic sites are not renewed because of the cost reduction of local governments. Therefore, some walkers who want to acquire culture complain about it.

This time, we made a system to get voice guidance of the historic sites from cell phones in consideration of the middle-aged walkers.

1.2 Purpose of Research

From a statistical data, we studied trends of cell phone users and the cellular technique in the future. Moreover, we picked out the technical and management problems of the trial system from a questionnaire. In this paper, we describe the effectiveness of the system.

2 Cell Phone Users

The history of cell phones dates to a car telephone service which was developed by Dendenkosha of the time in 1979. In 1985, shoulder phones were released with the liberalization of telecommunications. Since 1990s, telephones have been miniaturized and the cost has become cheaper. Then, with the digitization of communications network, i-mode models which had an Internet connection were released. The appearance of i-mode caused the explosive increase of customers.

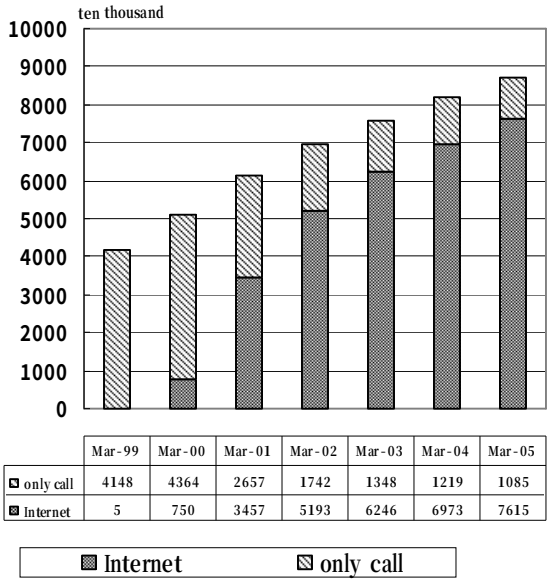
Figure 1 shows the transition of the number of cell phone contract from 1999 to 2005. It was increasing year by year. In 2005 the number was over 87 million and increased double during six years[2].

The number of Internet-connected cell phone contract reached 80% of the total in 2005. Cell phones are changing from just a tool of conversation to a tool of Internet connection.

Figure 2 shows the result of a survey against users what kind of functions their cell phones have.

The camera function is 84.4% with the highest percentage and the moving picture function is 63.8%. Then the Appli[3], the infrared communication function, the two-dimensional bar code reader and the function as a music player are following in that order[4]. However, in this question, users were allowed multiple answers and this result came from such a survey. Cell phones on sale in recent years have at least top 6 functions per cell phone. That is

the functions of a camera, moving picture, Appli, infrared communication, two-dimensional bar code reader and music player.



	Mar-99	Mar-00	Mar-01	Mar-02	Mar-03	Mar-04	Mar-05
only call	4148	4364	2657	1742	1348	1219	1085
Internet	5	750	3457	5193	6246	6973	7615

Figure 1: Transition of the Spread of Cell Phones
Ministry of Internal Affairs and Communications, Japan, "Information and Communications in Japan", White paper 2005, p.85

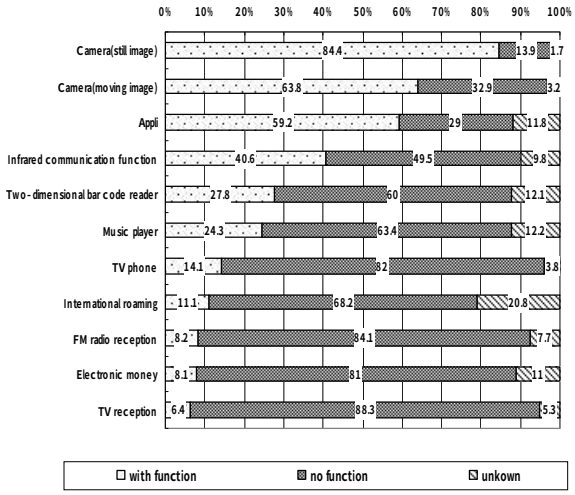


Figure 2: Presence of Each Function
Ministry of Internal Affairs and Communications, Japan, "Information and Communications in Japan", White paper 2005, p.87

Figure 3 shows the intention of users to utilize each function of cell phones.

After checking a research what function people want to use from now, including the existing functions, the camera function accounted for 72.7% with the highest percentage. The moving picture was 46.9%, the function as a music player was 35.8%, the Appli was 31.7%, and the function as a television was 27.6%. The number of cell phones with the function as a music player and a television is fewer compared with that of cell phones with the function as a camera and Appli. However, such cell phones which can be used as a music player and a television are very popular among young people. In the future, it is expected that cell phones with such functions will be developed widely[5].

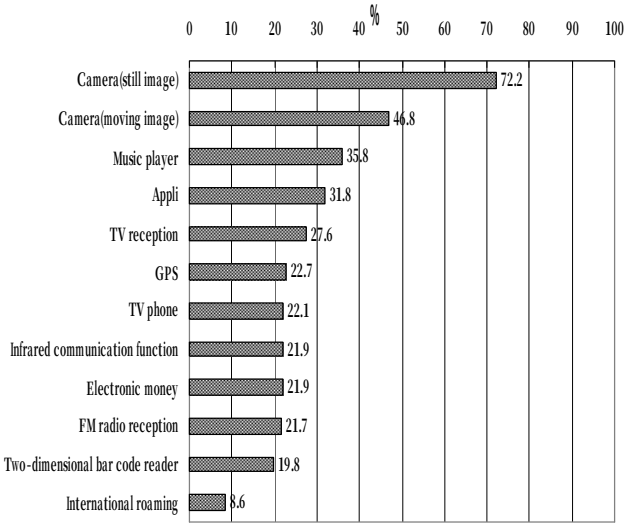


Figure 3: Intention to Use Each Function
Ministry of Internal Affairs and Communications, Japan, "Information and Communications in Japan", White paper 2005, p.87

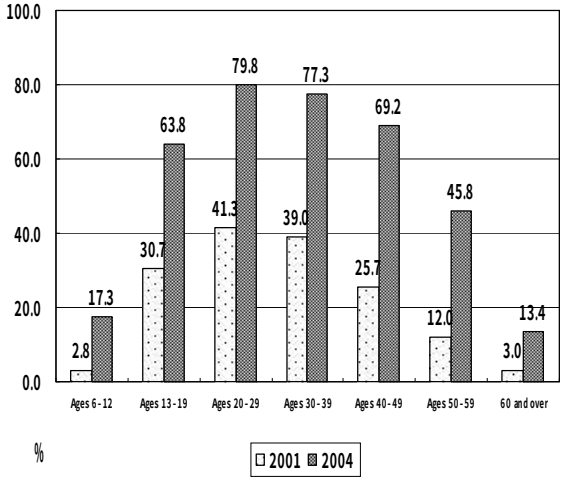


Figure 4: Rate of Internet Use with Cell Phones
Ministry of Internal Affairs and Communications, Japan, "Information and Communications in Japan", White paper 2005, p.118

Figure 4 shows the rate of Internet use with cell phones. Compared the rate in 2001 with the rate in 2004, the use of the Internet through cell phones has increased in every age group. The use rate between the ages of 13 to 49 accounts for over 60% but the

rate among users over the age of 50 has decreased on the contrary. There is still a generation gap. However, as time goes by, the disparity will be narrowed. Also, there are differences in other factors, gender, urban size and annual income. All things considered, the factor of age group has the greatest effect.

From the results stated above, users of the Internet through cell phones have been increasing steadily, and people are looking for value-added cell phones. That is a reason we say this is the time of ubiquitous computing.

3 Abstract of Experimental System

In this chapter, we clarify technical issues and look at problems of system administration and management on the basis of system configuration and a questionnaire. Then we study the usefulness of this system.

3.1 System Configuration

We made three systems on trial which could be used through NTT Docomo, Ezweb and Vodafone. Table 1 shows a data format for each carrier[6].

Table 1: Data format of Each Carrier

Carrier	Language	Multimedia	Graphics
NTT Docomo	C-HTML.,X-HTML	Mfi,3GPP	GIF,JPEG
Ezweb	HTML,X-HTML,HDML	SMAF,AMC,3GPP2	PNG,GIF,JPG
Vodafone	HTML,X-HTML	SMAF,3GPP	PNG,JPG

The languages of Table 1 are made on the basis of HTML language to built website and they are the specific standards of each carrier. Additionally, music file like MP3 is not played and set as ring alert on cell phones. Therefore, the reproduced sound using video files is substituted. As a result, 3GPP and 3GPP2 of multimedia can be played on Quick Time Player. Mfi and AMC are the specific standards of each carrier. Moreover, SMAF is the compression format of music which was developed by YAMAHA CORPORATION.

A trial system configuration diagram is illustrated in Figure 5 and the explanation is below.

The system developed here, converts the data with a text into the synthesized voice data, and compresses it. The system stores the compressed voice data in the database. And, the data can be accessible to the database from the cellular phone. The compressed voice data is downloaded by the XML program, and reproduced on the cellular phone. The following steps are performed.

(1) Description about each historic site is converted to text.

(2) (1) is converted to sound files with voice synthesis engine.

The voice synthesis engine used Via Voice of International Business Machines Corporation. The Via Voice is a software of the transformation from the input voice or text, to the output synthesized voice as the user interface, which is applicable to the conversation process between human and the computer.

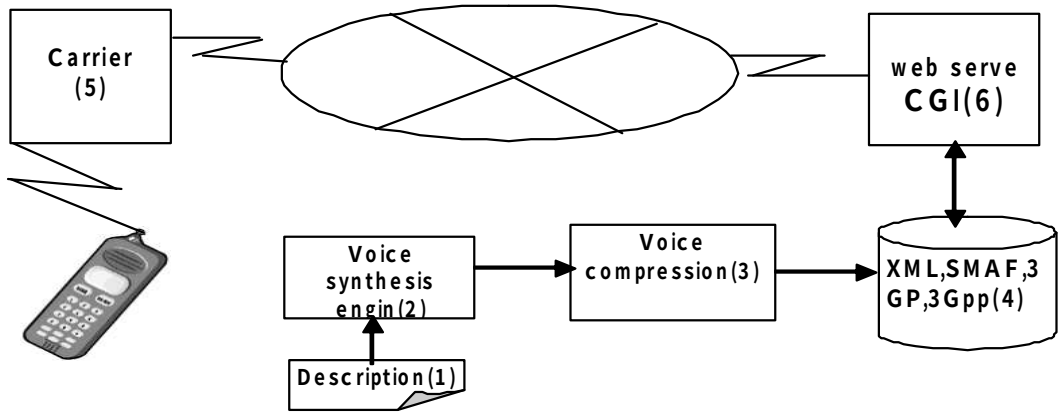


Figure 5: Diagram of System Configuration

(3) (2) is compressed into SMAF, 3GPP and 3GPP2 by carrier.

Human can recognize only the sound of about 20Hz～20 kHz as MP3, in principle.

Human do not hear any frequency, with the same sensitivity, but roughly 3～4 kHz with the top sensitivity. Then, the sensitivity of a low region and a high region is reduced. In a word, there is a limit in hearing the frequency of the sound, and the sound below the limit is not heard. The principle of cutting the sound is used as that of the compression.

(4) The voice data is downloaded from XML program and (3) is played on a cell phone.

When data with large volume was downloaded, it is necessary to have the specialized software in using the conventional HTML language. However, in the XML language, which was standardized by W3C, the download became possible easily on the Web server by specifying the data volume, the data name, and the volume type of data with using the object tag. Then, it is necessary to consider the interface with the data base.

Figure 6 shows an example of XML program on the screen.

```
<?xml version="1.0" encoding="Shift_JIS"?>
<!DOCTYPE html PUBLIC "-//OPENWAVE//DTD XHTML
1.0//EN" "http://www.openwave.com/DTD/xhtml-basic.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"
xml:lang="ja" lang="ja">
<html>
<head>

<td><center><object
data="http://www.geocities.jp/nari_pc/okazakijyou.3g2"
type="audio/3gpp2" copyright="no" standby=" Explanation 1
( click)">
<param name="disposition" value="devmpzz"
valuetype="data"/>
<param name="size" value="181775" valuetype="data"/>
<param name="title" value="okazakijyou"
valuetype="data"/></center>
```

Figure 6: XML Program

(5) The cell phone is accessed to a server from a carrier through the Internet.

(6) There are peculiar character strings that can be identified in OS of each cellular phone and the personal computer, respectively. This is called environment variables. The CGI programs are accessed from each cellular phone in the Perl language. Then, the access can be distributed according to the career by displaying the environment variables.

Figure 7 shows an example of the program.

```
#!/usr/bin/perl

$agent = $ENV{'HTTP_USER_AGENT'};
$str0 = "http://www.geocities.jp/nari_pc/ntt/okazaki.html";
if ($agent =~ "DoCoMo"){ $str0 =
"http://www.geocities.jp/nari_pc/ntt/okazaki.html"; }

"http://www.geocities.jp/nari_pc/okazaki.html"; }
   }
#print "Content-type:text/html¥n¥n";
print "Location: $str0¥n¥n";
```

Figure 7: Example of Perl Program

NTT Docomo use TCP/IP in protocol and Ezweb use UDP/IP in wireless communication. Each carrier has the advantages, such as the reliability and the high speed of data transmission, and also has the disadvantages. In recent cell phones, XML, proposed as global standard by W3C (The World Wide Web Consortium) on December 2000, is set as standard. Moreover, we can display still images using JPEG in cell phones of all carriers. In the meanwhile, the data format of multimedia and the program to download data differ from each carrier because of its protocol.

Vodafone and NTT Docomo allowed to use 3GPP as sound file format. However, we couldn't play sound files automatically in the Vodafone model, therefore, we changed the format from 3GPP to SMAF. Unfortunately, the capacity to convert to SMAF was limited.

In Perl program which is sorted by carriers, we adopted that each carrier could have special characters in environment variables.

Figure 8 shows a diagram of system contents. The main page is linked to the pages of historic sites. We added QR code to enable to use two-dimensional bar code.

The figure shows (1)Main screen, (2)Okazaki Castle, (3)Ooka Jinya, (4)Daijyuji Temple, (5)Deva Gate of Takisanji Temple and (6)Toshogu Shrine of Takisanji Temple. For instance, we take Description 1 of (2) below.

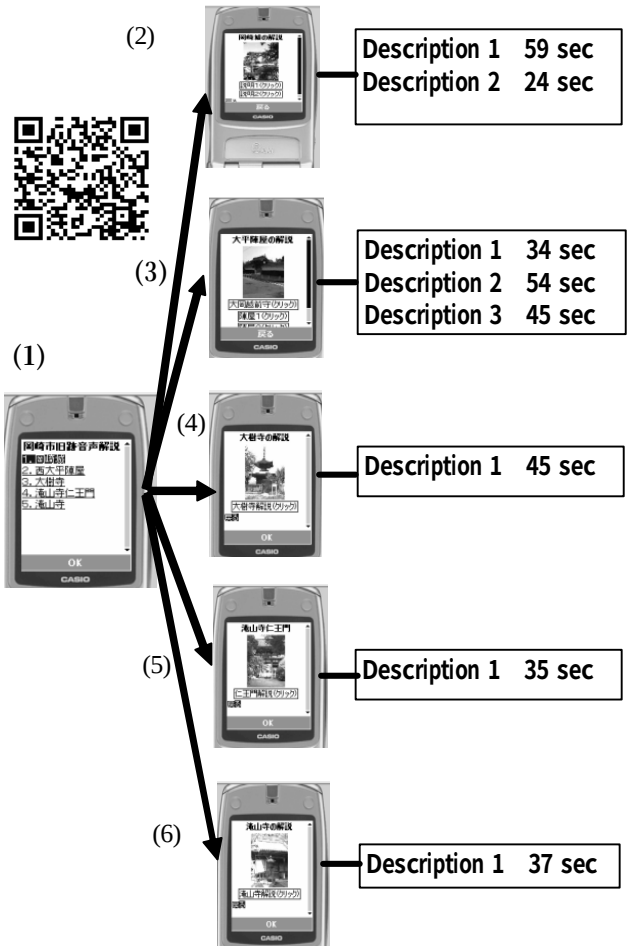


Figure 8: Diagram of System Contents

Description 1 of (2) (59 seconds) : "In the middle of the 15th century (Muromachi period), Saigo Danjyozaemon Yoritsugu built a castle in the place where the Okazaki Castle stands now at first. And then, Matsudaira Kiyoyasu made a triumphal entry and established the full-blown Okazaki Castle. Tokugawa Ieyasu was born in this Okazaki Castle on December 26, 1542 (Tenmon 11). In Edo period, Okazaki city was scared as 'The castle of promotion to God-King', and successive feudal lords, the Honda clan (descended from Yasushige), the Mizuno clan, the Matsudaira clan and the Honda clan (descended from Tadakatsu) became lords of the castle. It is said that feudal loads took pride in becoming a lord of this Okazaki Castle despite the yield was only 5 grain screen.€ 35

The reason why we used hiragana and brackets much in the description is that the sentence has words specific to reading and the voice synthesis engine (IBM viavoice v8) can not distinguish the words.

3.2 Evaluation of the System

We conducted a questionnaire on two groups, Group A (32 people in their 10s to 20s) and Group B (7 people in their 40s to 50s). The research of Group A was taken in doors and on the other hand, the research of Group B was gathered by going for walking actually. The number of subjects was small, because the respondents to the survey were limited to users who had specific cell phones. The evaluation of the questionnaire is 5-point scale. We show the survey contents below.

(1) Was the operability to output voice guidance good?
(2) Was the time to download voice data short?
(3) Was the description good?
(4) Was the voice clear? (sound quality)
(5) Was it economical to download?

Figure 9 shows the frequency distribution of Group A and Figure 10 shows that of Group B on the questionnaire.

Also, Table 2 shows the standard deviation of each group

It was normal distribution in Group A, but there were large variations and response bias in Group B. About the operability of Question (1), the average evaluation was 3.1 in both Groups A and B.

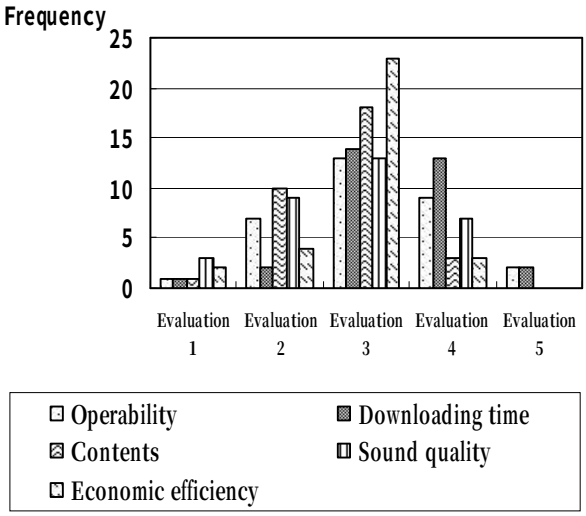


Figure 9: Frequency Distribution of Group A

If the program is devised we could receive higher reputation. About the download time of Question 2, the younger subjects in Group A felt slow because they often downloaded games or other programs, thus, the average evaluation was 3.4. In Group B, the subjects had little experience to download, thus, the evaluation was 4.6 on average. Carriers already have a plan of high-speed specifications, therefore, this problem is solvable.

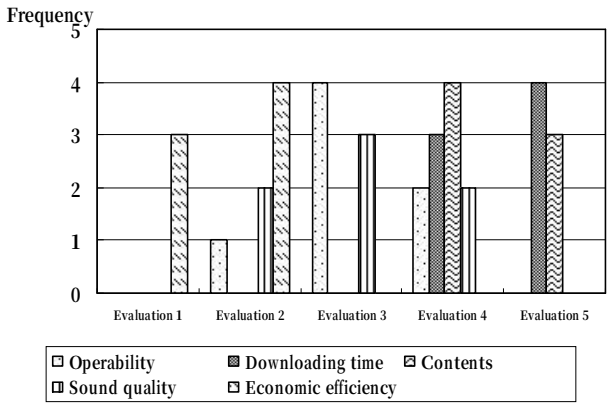


Figure 10: Frequency Distribution of Group B

Table 2: Standard Deviation

	(1)	(2)	(3)	(4)	(5)
Group A	0.164	0.146	0.119	0.159	0.118
Group B	0.241	0.187	0.187	0.286	0.187

About the contents of Question (3), the average evaluation was 2.7 in Group A. It is believed as the fact that the contents didn't attract young people's interest and the research was carried out in doors. Group B had an average of 4.4. This was the first time they used the system, therefore, the evaluation was high. However we should not use the same content. We need to update the information. About the sound quality of Question (4), the people in Group A, who were used to hearing music, made a strict judgment. For that reason, the evaluation was 2.8 on average. In Group B, the average evaluation was 3.0. Machine sound had a bad reputation. Though there is a possibility to ask an announcer to speak, if people get used to the machine sound, the problem would be solved. About the economic efficiency of Question (5), the average evaluation was 2.8 in Group A. The number of people who signed up for the flat-rate system was almost half, and the figure was lower than expected. Group B also gave low evaluation of 1.6 on average. Only 1 out of 7 subjects had an agreement on the flat-rate system.

The amount of the largest data is 229KB in this system. In case someone download the data on the assumption that one packet is 128B and 0.2 *yen*, it costs about 229B×1024/128B×0.2 *yen*=367 *yen* in disregard of the header.

In short, updating contents and economic efficiency are remaining problems.

4 Conclusions

In art galleries or museums, visitors can borrow a receiver to hear voice guidance with extra charge. We guess it costs a lot of money to manage and renew these devices. If customers download these contents to their cell phones or portable music players, the administrator could save much money. Though in an experimental stage, some zoos and theme parks plant an IC tag in the panel and provide sound explanation by using an exclusive receiver. About navigation system, the special markers are transmitted at every point and visitors can get positional information. Therefore, it is also complicated to manage this effort and the management costs money. However, if the management takes advantage of cell phones which 70% of the population has, it is possible to put such systems to practical use.

If hot spots are set in places desired, we can operate the system free of communication charge through customer's cell phone. IP phones to use the hot spots have already sold for companies.

The entry of many companies into every carrier will make the packet communication fees cheaper. We expect the development of techniques in each carrier.

References

[1] Ministry of Internal Affairs and Communications, Japan, "Information and Communications in Japan", White paper 2005, p.85

[2] Ministry of Internal Affairs and Communications, Japan, "Information and Communications in Japan", White paper 2005, p.85

[3] Application software

[4] Ministry of Internal Affairs and Communications, Japan, "Information and Communications in Japan", White paper 2005, p.87

[5] Ministry of Internal Affairs and Communications, Japan, "Information and Communications in Japan", White paper 2005, p.87

[6] Cell phone provider (NTT Docomo, AU and Vodafone in general)