STUDY ON REFERENCE MODELS FOR HMI IN VOICE TELEMATICS TO MEET DRIVER’S MIND DISTRACTION

Makoto Shioya, Senior Researcher
Systems Development Laboratory, Hitachi, Ltd.
1099 Ohzenji, Asao-ku, Kawasaki-shi, Kanagawa 215-0013, Japan
Tel: +81-44-959-0286, Fax: +81-44-959-0851
E-mail: shioya@sdl.hitachi.co.jp, Web: http://www.sdl.hitachi.co.jp/

Takuya Nishimoto, Research Associate
Graduate School of Information Science and Technology, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
Tel: +81-3-5841-6902, Fax: +81-3-5841-6902
E-mail: nishi@hil.t.u-tokyo.ac.jp, Web: http://hil.t.u-tokyo.ac.jp/

Juhei Takahashi, Senior Chief Researcher
ITS Center Research & Planning, Japan Automobile Research Institute (JARI)
1-30 Shiba-daimon 1-chome, Minato-ku, Tokyo 105-0012, Japan
Tel: +81-3-5733-7924, Fax: +81-3-5473-0655
E-mail: fjakahas@jari.or.jp, Web: http://www.jari.or.jp/

Hideharu Daigo, Senior Researcher
ITS Center Research & Planning, Japan Automobile Research Institute (JARI)
1-30 Shiba-daimon 1-chome, Minato-ku, Tokyo 105-0012, Japan
Tel: +81-3-5733-7924, Fax: +81-3-5473-0655
E-mail: hdaig@jari.or.jp, Web: http://www.jari.or.jp/

SUMMARY

This paper describes a study on the reference models for human machine interaction (interface) (HMI) in a voice telematics services. The goal of the reference models is to reduce the possibilities of drivers’ mind distractions even in the cases of hands-free and/or eyes-free driving situations. One of the basic principles is ‘Voice telematics services should be controlled so that driving workload plus telematics workload should not exceed driver’s ability.’ In the first step of the study, the actual driving tests were conducted in which HMI dialogues of the test telematics services were controlled based on emergency and priority considerations. The tests included right-hand turns as the typical situations where the driver’s workload increased. As a result of the tests, the necessity of the reference models was confirmed. Now, as the second step, we are drafting the reference models systematically based on the tests.
INTRODUCTION

BACKGROUND

A voice-activated human-machine interface (HMI) is an effective way of reducing the driver’s workload and visual distraction when operating an onboard information system to access the Internet and obtain telematics services.

However, it has been pointed out that even hands-free and voice-activated HMI technology may not always be safe to use depending on the nature of the service involved or the driving state of the vehicle at the time. This is believed to be caused by the mind distraction compared to the visual distraction.

CONCEPT OF VOICE-ACTIVATED SYSTEM

To study the mind distraction, we have constructed the architecture of a voice-activated system that incorporates network-distributed processing capability (Figure 1).

![Concept of Network-Distributed Voice-Activated Telematics Service System](image)

The architecture adopts dialogue control system (HMI server) which takes into account driving situation, emergency and priority to reduce mind distraction (Figure 2). The HMI server constantly monitors the situation in and around the vehicle, the driver’s state, the state of an HMI dialogue and external information with high priority and so on. If mind distraction tendency or high priority external information is sensed, the HMI server changes the nature of the dialogue according to the circumstances, if necessary. These are the distraction sensing function and the dialogue control function.

The tests have been undertaken to verify the effectiveness of the dialogue control system architecture.
DRIVING TESTS FOR CONFIRMING NECESSITY OF HMI REFERENCE MODELS FOR VOICE TELEMATICS

WORKLOADS, DRIVER’S SKILL & REMAINED ABILITY AND DIALOGUE CONTROL CONCEPT

The necessity of HMI reference models has been confirmed by the tests on the streets. The tests have been conducted in situations where the workload of obtaining a telematics service have been added to the workload of operating a vehicle. The relations between the workloads, two driver’s skills and the driving ability remained to each driver are shown in the Figure 3. The telematics load is assumed to be constant and as examples of the increased driving workload situations, passing, turning right or left and merging with traffic among others are shown in the figure.

Fig.3 An Example of Dialogue Control for Driver B
The control of HMI dialogues has been executed so that the total workload has not exceeded the driving abilities of individual drivers. An example of the HMI dialogue control for the driver B is shown in the Figure 3. In this control case, the dialogues are paused when the driving load plus the telematics load is assumed to exceed the driver B’s skill. The passing and the mixing cases correspond to this condition in the figure.

**DRIVING TEST SITUATION AND DIALOGUE CONTROL**

The main test situation is the right or left turn case. When a vehicle approaches an intersection and a driver switches the winker on to turn the intersection right or left, the HMI server senses the increase of the workload for the driver and has the proceeding dialogue to pause. When the vehicle turns and the driver switches the winker off, the HMI server resumes the dialogue paused. These processes are illustrated in the Figure 4 in a right turn case.

![Diagram of Dialogue Control in Case of Right Turn](image)

**Fig.4  Dialogue Control in Case of Right Turn**

A dialogue is paused in a turning period basically, but the figure shows a case when the dialogue is resumed before the end of the vehicle turn. This is the case when the vehicle speed is zero and the driver burges in and uses the command ‘Resume’. This function is prepared for realizing flexibility of the HMI control to adapt various situations and individual drivers.

**EVALUATIONS BY QUESTIONNAIRE SURVEYS**

After the tests, we have the evaluations of the tests shown below by questionnaire surveys to ten drivers.

**MAIN RESULTS OF THE QUESTIONNAIRE**
PRELIMINARY STUDY ON HMI REFERENCE MODELS FOR VOICE TELEMATICS

A preliminary study was made on the following items that should be shared in common so that a driver could dialogue with the system simply and safely while driving, regardless of the
vehicle types.

**STUDY ON VOICE COMMAND REFERENCE MODEL**

Voice commands are to be selected, assuming that telematics services are obtained within a framework that does not interfere with the execution of driving operations. In other words, sub task ( service access ) must not interfere with main task ( driving ). From this viewpoint, we selected two types of command functions:

(1) Command function that enables drivers to control the progress of a dialogue as they wish. Examples of commands are:
   - Start (Open): start dialogue
   - Wait (Pause): pause dialogue
   - Continue (Resume): resume dialogue
   - Jump: jump to coming service step
   - Back: back to past service step
   - Slowly: say prompt more slowly
   - Fast: say prompt more fast
   - Stop (Close): stop dialogue

(2) Command function for adjusting the degree of control over the progress of a dialogue. Examples of commands are:
   - Many, More
   - a Little, Little by Little
   - Reverse
These are used for quick decision of resume position and used together with ‘Back’, ‘Jump’, etc..
These are also used for adjusting speed of dialogue progress and used together with ‘Wait’, ‘Slowly’, ‘Fast’, etc..

**STUDY ON EFFECT SOUND REFERENCE MODEL**

Sound prompts or effect sounds are useful, assuming that a sound prompt, music and a pause will be needed to convey the state of a dialogue ( i.e., mode ) to drivers without involving unnecessary attention on their part in the process of obtaining a telematics service. In other words, sub task ( service access ) must not give drivers excessive load. Effect sounds, music, and timing are useful to inform progress to drivers with little load. From this viewpoint, we considered two types of effect sounds:

(1) Effect sounds for reassuring drivers by letting them know the status of the system. Examples of system status are:
   - Request Accepted: system accepted driver’s request
   - Under Processing: system is processing services ( searching, etc. )
     - No Waiting: get result soon
   - Waiting: take more time ( proceed to the next step is expected ).
     - processing service, reconnecting center.

(2) Effect sounds for letting drivers know what action they should take and the reason.
Examples of drivers’ action are:
- Utterance is Prompted: driver is prompted to utter (to say)
- System is Waiting Input for Prompt
- Repeat: system could not recognize
- Other Word: system recognized, but not expected word
- Malfunction: action such as reset, etc. is requested (waiting comes to nothing)
- Server is Malfunctioning
- Network Connection is Cut

STUDY ON HMI DIALOGUE REFERENCE MODEL

HMI dialogue control is necessary to be executed, according to safety, emergency and priority considerations and the driving skill of individual drivers. We considered four basic cases of controlling dialogues:

(1) Cases requiring control of a dialogue before a dangerous situation develops.
In these cases, the dialogues are paused, effect sounds are inserted and announcements with the reason are inserted.
For example, when winker is turned on (in right or left turn, passing, etc.), the system informs kindly to the driver and controls the dialogue carefully.

(2) Cases where a potentially dangerous situation has developed.
In these cases, the dialogues are paused quickly, and no effect sound or announcement.
For example, in sudden braking, effect sounds or announcements are too much for the driver.

(3) Cases where the vehicle receives high priority information from an external source, such as information concerning traffic restrictions.
In these cases, the dialogues just in progress are paused and the dialogues concerning the external high priority information are started.
For example, the system informs the driver before the right turn, if the street after the right turn is impassable.

(4) Cases where a danger warning is issued by a system that monitors safe vehicle operation.
In these cases, the dialogues are paused quickly and give priority to the safe vehicle operation monitoring system.
For an example, give top priority to hazard alarm from a collision alarm system.

RELATED HMI ISO STANDARDS AND THEIR RELATIONS WITH OUR REFERENCE MODEL STUDIES

There are some ISO standards in ITS area. Typical ones from the ISO in-vehicle HMI related standards are shown in the Figure 5.

Two of the standards are in public, we briefly outline their relations with our reference models for voice telematics studied above.

(1) ISO 14813-1 Reference Model Architecture for the TICS Sector - Tics Fundamental Services (2)
This standard defines 32 ITS (TICS) services. Services are classified in 8 categories. They are traveller information, traffic management, vehicle, commercial vehicle, public transport, emergency, electronic toll/fee collection, and, safety.

If some of the telematics services, such as e-mail, information retrieval, are defined in addition, it is sufficient for the definition of the telematics services.

**ISO 14813-1 Reference Model Architecture for the TICS Sector**
- TICS Fundamental Services

**16951 Road vehicles**
- Ergonomic aspects of transport information and control systems (TICS)
- Procedures for determining priority of on-board messages presented to drivers

**ISO 15005 Road vehicles**
- Ergonomic aspects of transport information and control systems
- Dialogue management principles and compliance procedures

**15007 Road vehicles**
- Measurement of driver visual behaviour with respect to transport information and control systems

**15008 Road Vehicles**
- Ergonomic aspects of transport information and control systems
- Specifications and compliance procedures for in-vehicle visual presentation

**15006 Road Vehicles**
- Ergonomic aspects of transport information and control systems
- Specifications and compliance procedures for in-vehicle auditory presentation

**17287 Road Vehicles**
- Ergonomic aspects of transport information and control systems
- Procedures for assessing suitability for use while driving

**Fig.5 Typical HMI Related ISO Standards**

(2) ISO 15005 Road vehicles - Ergonomic aspects of transport information and control systems - Dialogue management principles and compliance procedures (3)

This standard presents ergonomic principles for the design of the dialogues between the driver of a vehicle and the ITS systems while the vehicle is in motion. The standard aims to reduce
driver’s workload as a whole for safe and effective control of the vehicle.

The important principles of an ITS dialogue are appropriateness for use while driving (compatibility with driving; simplicity; timing / priorities), for the ITS task (consistency; controllability), and for the driver (self-descriptiveness; conformity with driver expectations; error tolerance).

Here, the dialogues are assumed to be performed mainly by tactile input, visual display and auditory display. This dialogue management standard is very close to our concern with the dialogue control to reduce mind distractions mentioned in the ‘INTRODUCTION’ section. If we add some special descriptions concerning dialogues using voice recognition as an input, it will become sufficient standard for the voice telematics. The descriptions may be principles, requirements, recommendations, and examples.

**HMI REFERENCE MODEL FOR VOICE TELEMATICS**

Taking the preliminary study on HMI reference models into account, we are now drafting the reference model for voice telematics. In this section, we describe first, the viewpoints of the reference model scope and basic principles and some requirements for the design of the dialogues between the driver and the voice telematics service systems in a vehicle. Next, we describe the evaluations of our activities by questionnaire surveys to auto makers and car navigation vendors in Japan.

**HMI REFERENCE MODEL STUDY FOR VOICE TELEMATICS**

**REFERENCE MODEL SCOPE OF THIS STUDY**

We describe here the viewpoints of our HMI reference model scope for voice telematics.

Telematics services are based on the rapidly developing information and telecommunication technologies. They include not only traffic information, POI (point of interest) information, news, and weather, but also various services, such as e-mail, electronic commerce, information retrieval, and entertainment, which are popular in the Internet world. Because of wide spread information categories and their complicated information hierarchies, accessing procedure may become complicated so that accessing through the sequence of tactile inputs becomes hard from the viewpoint of hands-free and eyes-free requirements while the vehicle is in motion.

Instead, applying voice input techniques to the telematics service access is suitable as for the hands-free and eyes-free requirements. But it is needed proper consideration for the dialogues between the driver and the telematics service systems even in voice input adopting case, for ensuring safety. This is because the dialogues using voice input may become not simple, so that the mental workload increases to the critical level for safe driving according to the driver’s condition and/or the other driving situations. This mental workload must be decreased so as the mind distraction of the driver not to occur.

The viewpoints of our reference model scope come from the above consideration. Those are:
Concerning the dialogues between a driver and telematics service systems while the vehicle is in motion.

The above dialogues are executed mainly using voices as inputs by the driver. These voice inputs are recognized mainly using voice recognition techniques.

The above telematics service systems treat services not only traffic information, POI information, news, and weather, but also various services, such as e-mail, electronic commerce, information retrieval, and entertainment, which are popular in the Internet world.

The reference model aims to present requirements and recommendations for accessing the telematics services within the driver’s mind-free limits.

The reference model suggests basic principle, minimum requirements and recommendations for the design of the dialogues, using voice primarily, between the driver and the voice-activated telematics service systems while driving. Most suggestions in the reference model are included already in the ISO standards described above, so that the reference model aims to supplement and enforce them from the standpoint of dialogues using voice which realize hands-free and eyes-free states. In addition, the reference model aims to suggest the minimum items that make drivers access telematics services within mind-free states.

**Basic Principle and Requirements of Reference Model**

We describe here the basic principle and some basic requirements for the dialogues of the telematics services.

(1) Basic Principle
In the dialogues for the telematics services, the important basic principle is ‘mind-free’ state of a driver. The mind-free state could occur under the condition that the driver’s skill is greater than the driving load plus telematics load.

(2) Basic Requirements:
- The dialogues for the telematics services are to be managed or controlled so as to the sum of the driving workload and the telematics workload does not exceed the driver’s workload processing skill or capability as a whole.
- The driving workload and the telematics workload are to be pre-estimated or measured real-time.
- The driver’s skill or capability is to be pre-estimated or measured real-time.
- The management or control is to be done considering driver’s condition and the other various driving situations of that time.
- The management or control is to be done considering the priorities of the services.
- The telematics service systems’ state is to be informed to the driver properly.
- The final initiative of the dialogues is to be taken by the driver.

**The Next Study**

These principles and requirements will be broken down into details in the next study. Most of the details will correspond to existing principles, requirements, recommendations, or
examples of the ISO standards. A few will correspond to additional ones of the ISO standards. The reference model will be expected to supplement and enforce ISO standards from the standpoint of voice-activated dialogues within mind-free states.

**NECESSITY OF OUR ACTIVITIES**

We have made questionnaire surveys in order to make sure whether our activities of making draft reference model are in the right direction or not. The surveys were concerned about the standardization of voice-activated systems and addressed to the telematics service providers and car navigation vendors in Japan. The answers expected were those from the standpoints of researchers in voice technologies or engineers of car navigation technologies. We got answers from twelve professionals.

The main items and the results of the survey were as follows.

(1) Voice dialogue items to be standardized (multiple answers):

<table>
<thead>
<tr>
<th>Item</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>conditions of dialogue interruption</td>
<td>7</td>
<td>58%</td>
</tr>
<tr>
<td>conditions of dialogue resumption</td>
<td>3</td>
<td>25%</td>
</tr>
<tr>
<td>explanation of state</td>
<td>1</td>
<td>8%</td>
</tr>
<tr>
<td>voice commands</td>
<td>7</td>
<td>58%</td>
</tr>
<tr>
<td>effect sounds</td>
<td>4</td>
<td>33%</td>
</tr>
<tr>
<td>others</td>
<td>1</td>
<td>8%</td>
</tr>
<tr>
<td><strong>total</strong></td>
<td><strong>23</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

(2) Effect sounds to be standardized:

<table>
<thead>
<tr>
<th>Item</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>all effect sounds</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>basic effect sounds</td>
<td>6</td>
<td>50%</td>
</tr>
<tr>
<td>functions of effect sounds</td>
<td>3</td>
<td>25%</td>
</tr>
<tr>
<td>unnecessary</td>
<td>3</td>
<td>25%</td>
</tr>
<tr>
<td>others</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td><strong>total</strong></td>
<td><strong>12</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

(3) Voice commands to be standardized:

<table>
<thead>
<tr>
<th>Item</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>all commands</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>basic voice commands</td>
<td>8</td>
<td>67%</td>
</tr>
<tr>
<td>functions of voice commands</td>
<td>1</td>
<td>8%</td>
</tr>
<tr>
<td>unnecessary</td>
<td>3</td>
<td>25%</td>
</tr>
<tr>
<td>others</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td><strong>total</strong></td>
<td><strong>12</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

This results show more than half of the responders for the surveys feel the necessity of standardization in items of the dialogue interruption conditions and the voice commands. As
for the effect sounds and the voice commands, the basic sounds or voice commands have rather high standardization need.

Considering that these results are obtained not from the end user-side but from the maker-side, we are convinced of our activities of making reference model are in the right direction.

**CONCLUSIONS**

It reveals that mind-free concept is important even in voice-activated telematics services system, besides so called hands-free and eyes-free concept. We conclude that controlling the voice dialogues in case of high workload on drivers is effective for reducing mind distraction. It become clear that standardized voice commands and dialogue patterns increases usability of the system.

It is desirable for drivers to be able to obtain telematics services safely with HMIs configured on the basis of the same standard regardless of which car model they drive. To realize this, reference model and its basic principle and requirements are discussed. They are to be broken down into details in the next study.

This research and development project represents only the first step toward the construction of a system that makes effective use of voice activation for obtaining telematics services from the standpoint of safety. Toward that end, the authors work together with various organizations in conducting quantitative tests and in developing reference models with the aim of creating a common standard.

**ACKNOWLEDGEMENTS**

This paper is the result of the research by the Voice Telematics Working Group of JARI and of the preceding three years research and development project by the Association of Electronic Technology for Automobile Traffic and Driving (JSK: reorganized into JARI), under contract with MSSF, founded by METI. Our special thanks are due to Professor Tetsunori Kobayashi of Waseda University for guiding our work. We also express our thanks for the members of the Working Group and the related organizations and persons

**REFERENCES**

(1) Makoto Shioya, et alii, 10th ITS World Congress, Madrid, 2003 ‘Research on Network-Distributed Voice-Activated System Architecture for Telematics Services’
(2) ISO 14813-1 Reference Model Architecture for the TICS Sector - Tics Fundamental Services
(3) ISO 15005 Road vehicles - Ergonomic aspects of transport information and control systems - Dialogue management principles and compliance procedures