

DRIVER'S LOAD OF VOICE INTERACTION SYSTEM IN VEHICLES

Yonosuke Miki, Noritoshi Yoshitsugu, Kazuya Ito

Nissan Motor Co., Ltd.

Japan

Paper No. 170

ABSTRACT

On-board voice interaction systems such as voice activation system or text-to-speech (TTS) system enable drivers to operate devices or to obtain the desired information without relying on their visual processes. Although these systems are thought to reduce the driver's workload, few papers published papers have quantified any reduction. This paper describes the tests that were carried out to determine its potential influence on drivers' mental distraction. Tests were conducted with a driving simulator and with an actual vehicle. Subjects were given various on-board device tasks to operate while driving. Their eye motion, response time to the illumination of the LED installed in a vehicle and vehicle behavior data were recorded when the tasks were performed. The results obtained for all of the indices show that the mental distraction level when listening to a TTS reading of information is comparable to that of listening to the car radio and the workload of the voice activation system is significantly lower than that of a traditional manual operation system.

INTRODUCTION

Both voice recognition technology and TTS technology, which convert text information into synthesized speech, have made rapid progress in recent years. This progress has made it possible to develop on-board voice interaction systems, such as a voice activation system to control audio and navigation, or TTS synthesizer to speak whenever latest news is desired, or e-mail and other text data is downloaded from an external network via a cellular phone. At Nissan, we have developed a system that supports the provision of such services in Japan. Although an information device with TTS synthesizer might be expected to decrease the visual distraction, there remains the possibility that listening to information by the drives might cause a delay in judgment due to mental distraction. There are methods for quantifying distraction caused by visual information such as navigation map screens. One such method is the occlusion technique based on the glance time for operating equipment control switches. However, there are no effective indices at present for evaluating mental distraction induced by spoken information. Consequently, there is a need to investigate the effect of equipment usage on driver behavior by using some method other than driver observation, such as vehicle control performance or indices

derived from a model of driver recognition and judgment while operating a vehicle.

EXPERIMENTAL DESIGN

Basic concept of experiments

When in the process of driving, drivers are constantly repeating a cycle of information acquisition, recognition/judgment, and execution of driving actions or operating a vehicle. Acquisition of visual information is particularly important while driving. Obstruction of a driver's field of view can result in a change in vehicle behavior, increased anxiety or other undesirable effects. For that reason, in evaluating the operability of navigation systems and other devices for presenting visual information to drivers, an accepted technique is to quantify the length of time a driver glances away from the road (i.e., visual distraction time), which can be measured (1)(2). With in-vehicle services that combine speech recognition and TTS technology, it is assumed listening to spoken information or saying control commands may also affect driver behavior. However, it is difficult to quantify that effect based on external observations such as eye movement analysis. Therefore, the indices used in this study for evaluating mental distraction were determined on the basis of a driver recognition/judgment model while driving. The voice interaction task was evaluated for each index in comparison with various other on-board device tasks given to the subjects, such as listening to the radio, talking on a cellular phone, or viewing a navigation display, in order to examine the effect on mental distraction.

Evaluation indices (experimental measures)

A general model of driver recognition/judgment while operating vehicle is shown in Fig. 1. This model assumes that the driver's total processing capacity is allocated between that needed for driving, including confirmation of the external traffic environment, and that needed for operating in-vehicle devices and entering/receiving information. As indicated in condition (a) in the figure, even in situations with a high driving workload, such as when traveling in congested traffic or on a winding road, the driver has additional processing capacity provided that the workload involved in operating in-vehicle devices, or entering/receiving information, is low. In

contrast, under condition (c) where the operation of in-vehicle devices, or the entering/receiving of information, requires greater processing capacity, judgmental delay or error can occur due to the lack of any spare capacity. As a result, that can have the effect of influencing driver control over the vehicle, or the mental workload perceived by the driver.

Therefore, the mental workload required for voice-activated operations or listening to a TTS reading can be evaluated by measuring indices such as vehicle control or the driver's spare capacity. The following items were used in testing as the evaluation indices of vehicle control performance, added attention capacity and the overall mental workload. In addition to those indices driver's eye movement was recorded to measure glance time.

Vehicle control performance

This was measured by the standard deviation of the amount of change in the following distance to the preceding vehicle and the standard deviation of the host vehicle's lateral lane displacement.

Spare capacity

This was measured by the secondary task response time for pressing a steering wheel button when an LED mounted directly in front of the subject was illuminated while the person was operating an in-vehicle device or listening to/reading information.

Experiment II; to investigate the workload of voice activation system.

Tests were conducted on a driving simulator. The results obtained were then validated by conducting driving tests on a closed course with an actual vehicle.

Experimental conditions

Table 1 shows the experimental conditions designed for Experiment I.

Table 1.
Experimental Conditions

	Driving simulator tests	Actual vehicle tests
No. of test subjects	20 (15 men; 5 women)	6 (all men)
Age	25-40	25-42
Test course	Simulation of expressway driving	Oval proving ground course
Driving conditions	<ul style="list-style-type: none"> ● Following a preceding vehicle at 100 km/h ● Maintaining a constant following distance of about 40 m ● Presence of other vehicles 	<ul style="list-style-type: none"> ● Following a preceding vehicle at 60 km/h ● Maintaining a constant following distance of about 40 m ● Presence of other vehicles

Experiment II was conducted only on a driving simulator. The Eleven subjects were evaluated, and other conditions were same as driving simulator tests in Experiment I.

Tasks

Each subject was asked to perform two tasks simultaneously while operating either the driving simulator or the actual vehicle. As explained below, one task involved the operation of an in-vehicle information device and the other was a secondary task which was designed to measure driver's spare capacity

Information Device Tasks

In Experiment I, spoken information may be presented in a variety of forms. For instance, it may consist of long sentences, as in the case of e-mail. Alternatively, it may involve listening to information that can be partly anticipated such as weather reports. In view of these diverse forms, various listening tasks were designed for these tests as outlined in Table 2. In addition, entry of the destination in the navigation system, which involves a more complicated operating procedure, was also given as an operational task in order to make a comparison between auditory and visual tasks. The task of listening to a TTS reading also involved giving voice commands, i.e., the TTS reading began after a subject uttered the command three times. The subjects performed all of the tasks in the driving simulator tests, but from a safety standpoint, they did not perform the telephone and complicated navigation system tasks while driving the actual vehicle.

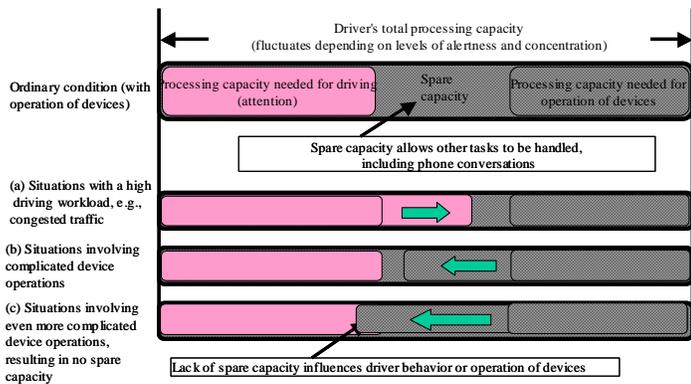


Fig. 1. Model of Driver Recognition and Judgment.

Subjective evaluation of mental workload

This was measured by the Adaptive Weighted Work Load (AWWL) calculated from the National Aeronautics and Space Administration Task Load Index (NASA-TLX) scores (3).

Methodology

Two experiments were conducted. The purpose of each experiment was as follows.

Experiment I; to investigate the workloads of information hearing via TTS system.

Table 2.
Information Device tasks

Task	Device	Task Time	Task details
Listening to the radio	Radio	35-46 sec.	Subject listens to a news recording played through the speakers. (no operation involved)
Listening to TTS reading	--	28 sec.	Subject presses the voice switch on the steering wheel, gives voice commands to call up spoken e-mails about meeting places or listens to news.
Reading text information	Navigation system display	1-2 sec.	Subject reads text information consisting of 20 kanji characters shown on a display at the top of the center instrument cluster.
Entering destination	Navigation system	20-30 sec.	Subject enters the destination by inputting a 10-digit phone number using a joystick.
Receiving a call	Cellular phone	5-8 sec.	When the cellular phone rings, subject retrieves it from the passenger's seat, puts it to his ear and confirms the sound.
Making a call	Cellular phone	5-8 sec.	Subject retrieves the cellular phone from the passenger's seat and calls the third number registered in the memory.
Answering a question	Cellular phone	60 sec.	Subject answers a simple question. Example: "Did you eat breakfast?" "Yes"
No task (control condition)	--	60 sec.	No operational task

In Experiment II, subjects were asked to operate conventional audio or navigation system manually. In addition, they were asked to operate voice activation audio or navigation system. Tasks were as follows:

- Input telephone numbers
- Search for station names
- Input addresses
- Increase track number of CD
- Select radio channels

Each task was consisted of a number of transactions varied from one to seven. Simple tasks such as audio control tasks consisted of a few transactions. Complicated tasks were designed to involve more than 6 transactions. Guidance of the operations and the recognition results of the command spoken by a driver were presented from voice activation system via voice by using TTS technology. Voice activation system can show the recognition results by text on the display. To investigate the influence of the display, some tasks were carried out without display.

Secondary task

The LED mounted at the top of the instrument panel was illuminated for one second at random intervals of 5-10 seconds while the subjects were performing the information device tasks. The subjects were instructed to push a button mounted on the right side of the steering wheel as soon as they noticed that the LED was illuminated. The configuration of the experimental setup is shown in Fig. 2. and Fig.3.

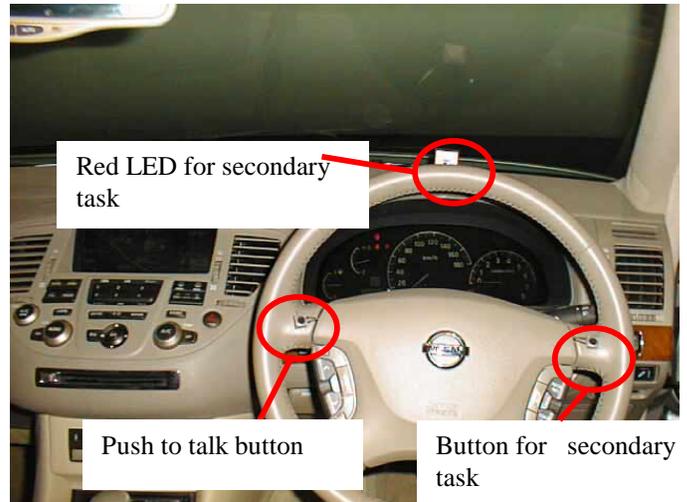


Fig. 2. Configuration of Experimental Setup.



Fig3. Driving simulator.

Experimental procedure

The tests were conducted according to the following procedure.

1. The subjects practiced the tasks outlined in Table 2 before the test began. They practiced to the extent that they were able to complete each task without making a mistake.
2. After practicing for about one minute to familiarize themselves with the operation of the driving simulator or the actual vehicle, the subjects then performed only the above-mentioned secondary task under a no-task

condition (control condition) and measurements were made for the evaluation indices.

- The subjects were asked to perform the secondary task simultaneously with the information device tasks in Table.2. They were informed in advance about the content of the radio broadcast and the TTS reading so that they would be able to follow the spoken information well.
- The tests were repeated again under the control condition and the evaluation indices were measured.

At the completion of each information device task, the subjects momentarily stopped the driving simulator or parked the vehicle, and filled out a NASA-TLX form.

RESULTS

Experiment I

Figures 4-7 show the results measured for the evaluation indices when the subjects performed the tasks in Experiment I. Similar tendencies were seen for the secondary task response times, lateral lane displacement and subjective mental workload ratings (NASA-TLX scores). Even when the subjects had listened to TTS reading, which combined entry of a voice command, or listened to the radio, their response times with respect to each evaluation index did not show any statistically significant differences when compared with the corresponding results for listening to the radio or for the control condition. However, the driver's manual operations for destination entry into the navigation system and visual concentration on the display screen produced significant differences ($P < 0.05$) compared with the control condition.

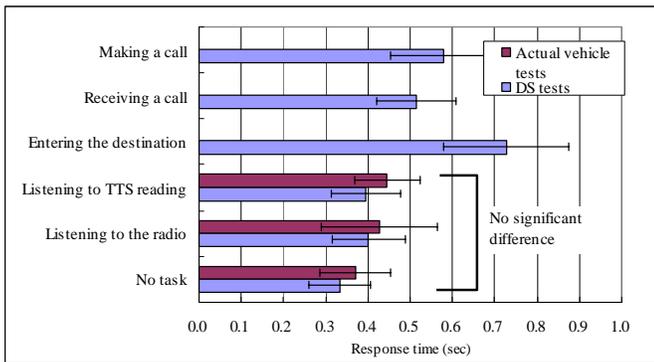


Fig. 4. Secondary task Response Time.

The standard deviation of the following distance (Fig. 5), however, showed a different tendency from those seen for the other evaluation indices. Only the task of entering the destination showed a significant difference ($P < 0.05$) compared with the other tasks. Moreover, a comparison of the response times in the driving simulator and in the actual vehicle (Fig. 4) showed that there were no statistically significant differences between the control condition, listening to the radio and listening to TTS reading

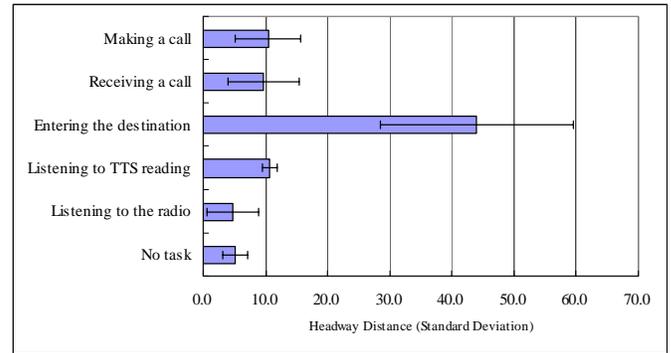


Fig. 5. Comparison of Variation in Following Distance (Standard Deviation).

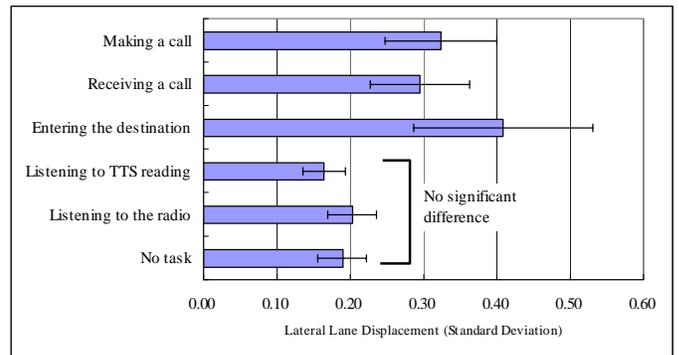


Fig. 6. Comparison of Lateral Lane Displacement (Standard Deviation).

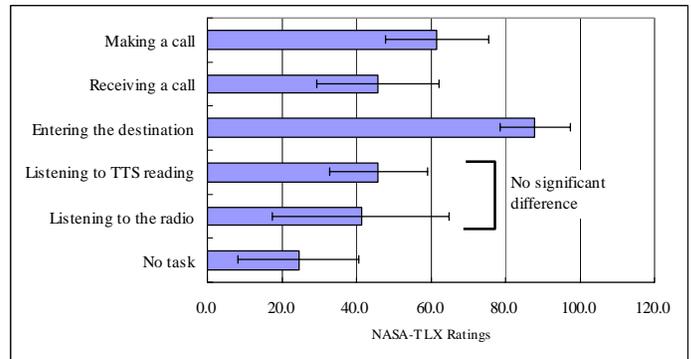


Fig. 7. Subjective Evaluation of Mental Workload.

Experiment II

The data shown in Fig.8 are the results of driver eye movement analysis when the subjects performed manual operations or voice activation tasks in Experiment I. TGT (Total Glance Time) of the voice activation tasks are significantly lower than those of manual operation tasks as expected. TGT of manual operation tasks increase linearly with number of transactions, whereas TGT of voice activation tasks remains low. Fig.9 shows the NASA-TLX Ratings in Experiment I. The ratings of voice activation tasks are lower than those of manual operation tasks. Even most complicated

voice task results are lower than most simple manual operation task.

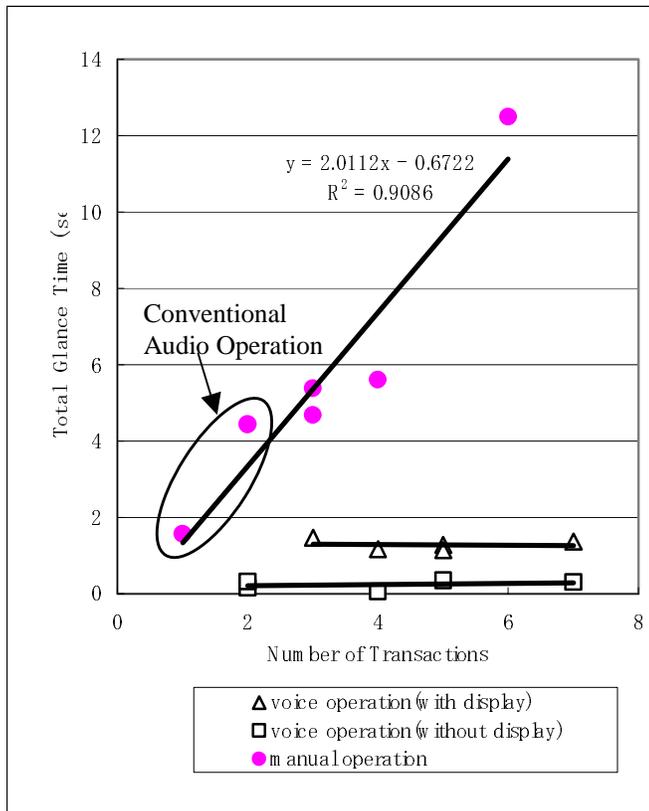


Fig. 8. Correlation Between Number of Transactions and Total Glance Time

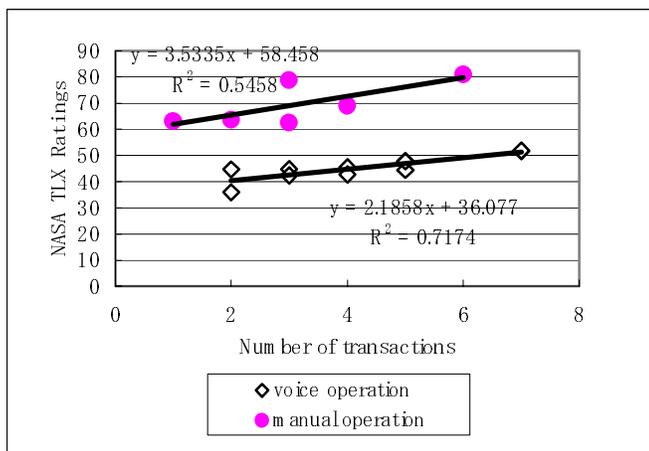


Fig. 9. Correlation Between Number of Transactions and NASA-TLX Ratings.

DISCUSSION

Voice interaction effect on driver behavior

For the task of listening to a TTS speech, the secondary task response time and lateral lane displacement did not show any significant difference compared with the results for listening to a radio. This suggests that the use of TTS technology to modify the acquisition of text information to an auditory task

would not have much influence on actual driving operations, although the act of listening by itself may increase the perceived mental workload. Even in the case of e-mail and the level of mental distraction would be comparable to that for listening to the car radio.

Because of less visual demand, mental workload of voice activation is significantly lower than that of manual operation. Even conventional manual audio operation tasks indicate higher mental workload value than complicated voice activation tasks.

Evaluation indices of voice activation

The result of following distance displayed a different tendency from the other indices. It showed that there were virtually no differences among the different tasks except entering destination task. This is thought to be attributable to the fact that the subjects paid the greatest attention to maintaining enough following distance to avoid a rear-end collision when they performed the various tasks. However, entering a destination in the navigation system required visual concentration on the screen when inputting information, making it difficult to maintain a constant following distance. In fact, the average distance for the other tasks was 35-50 m, whereas increased to 73 m for entry of the destination in the navigation system. This substantially longer distance would seem to indicate that the subjects subconsciously reduced their driving speed to prevent accidents when performing this complicated task.

The correlation of the secondary task response time and lateral lane displacement, which displayed similar tendencies as noted earlier, with the NASA-TLX scores is shown in Figs. 10 and 11, respectively. The results indicate that both of these indices correlated closely with the NASA-TLX ratings. Especially the reaction time value was highly correlated to the NASA-TLX scores. This result shows that the response time can be an effective evaluation index.

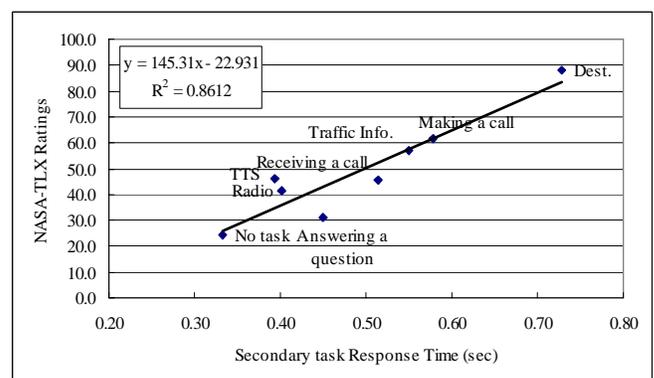


Fig. 10. Correlation Between Secondary task Response Time and NASA-TLX Ratings

However, if a test course involves winding roads, this button location could influence steering behavior itself, or if primary tasks are complicated, there is also a possibility that asking a subject to operate a button at the same time some other switch is operated could cause confusion, resulting in operational error.

considered in our ongoing research into methods for quantifying drivers' mental workload.

REFERENCES

[1] Kimura, K et. al. (1999). In-Vehicle Navigation System Operability while Driving. Proceeding of 6th World Congress on ITS.

[2] Niiya, K. (2000). ISO/TC 22/SC 13/WG 8N289 Trial Evaluation Results on Navigation Operation using Occlusion Technique. Discussion Paper on ISO/TC22/SC13/WG8 Visual Distraction TF.

[3] Miyake, S. Kumashiro, M. (1993). Subjective mental workload assessment technique – An introduction to NASA – TLX and SWAT and a proposal of simple scoring method -. The Japanese Journal of Ergonomics Vol. 29, No. 6

[4] E.R. Boer and A. Liu (Eds.) (1997). Cambridge Basic Research 1997 Annual Report. CBR TR 97-7, Cambridge Basic Research, Cambridge Massachusetts.

[5] Development of a Steering Entropy Method for Evaluating Driving Workload. Proceeding of 1999 JSAE Annual Congress.

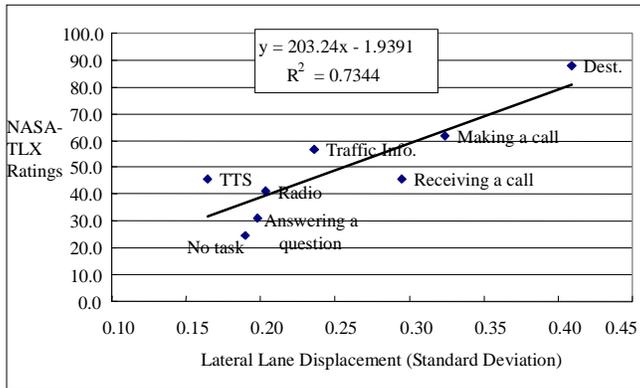


Fig. 11. Correlation Between Lateral Lane Displacement and NASA-TLX Ratings.

Accordingly, if the primary task is complicated, it is thought that lateral lane displacement can be an effective index for evaluating the driver's mental workload for tasks, because this index does not require any additional task and easy to measure if a driving simulator is used.

CONCLUSIONS

The following results were obtained in this study based on a comparison of the driving test data and the various types of quantitative data obtained from the subjects.

1. The use of TTS technology to transform text information acquisition into a listening task should not have a significant effect on driving behavior even in the case of information involving relatively long sentences. The mental workload involved in listening to spoken information is estimated to be similar to that of listening to the car radio.
2. The mental workload of voice activation system is significantly lower than that of manual operation systems. The mental workload involved in activating complicated tasks such as navigation operations by voice is estimated to be lower than that of operating audio manually.
3. Among the evaluation indices used in this study, the host vehicle's lateral lane displacement and the secondary task response time were found to be effective in estimating the mental workload involved in listening to spoken information and in operating equipment by voice commands. Lateral lane displacement in particular is thought to be highly useful because it is easy to measure with a driving simulator and does not require test subjects to perform any particular task.

With respect to lateral lane displacement, the steering entropy method have also been established for quantifying perturbations in steering smoothness caused by external factors aside from driving tasks (4)(5). The potential application of this technique and other methods will be