In critical situations in the railway system, people often have to rely on what is displayed on monitors without being able to see the real-life situation. Is it possible to support what is generally only shown as visual information by means of the other senses and thus increase safety? Today, audible information is sometimes used. New technologies offer haptic information. How well can visual information be combined with haptic information? Siemens’ Mobility Management Business Unit develops reliable operator console systems for mass transit and mainline operations throughout the world. Today, desktop PC systems with several monitors are used as operator console systems almost exclusively. As part of a research project, Siemens is developing a new solution for an operator touchscreen console with a reliable display on the basis of haptic and audible information. Instead of mouse clicking, the user interacts directly with his/her finger which, from the point of view of human factors, is a different type of interaction. This paper shows how a solution of using haptic and audible information for a reliable display can function in the railway system. The paper also describes two studies which have studied the effectiveness of this solution. The discussion below is based on the findings presented in Dose (2014), which have been further detailed for this publication.

**Introduction**

Smartphones and tablet PCs have become widespread and common over the last few years. Their proliferation has been largely confined to the personal sphere so far, but it appears likely that, thanks to their compact design and ease of use,
tablet PCs in particular begin to penetrate the commercial world as well – and with it the rail transport sector.

Tablet PCs, which are operated directly with fingers rather than via an intermediate mouse and/or keyboard like a conventional PC, open up new possibilities for user interaction via the sense of touch. Conventional PCs interact primarily visually, but a fingertip-operated tablet PC can also directly target the sense of touch. Since people have the ability to take in different information via different senses in parallel, this makes it possible to communicate multiple messages concurrently (Hampel, 2011). A tablet PC can transmit information not just visually and aurally, but also haptically.

The dependence on essentially one sense alone engendered by conventional PCs can be regarded as slightly unnatural, as people usually apply a multisensory approach to information gathering in the real world (Seebode, 2010). When in conversation, for example, people generally look at the other people involved as well as speaking and listening to them (Hoggan and Brewster, 2006). Machines, on the other hand, often rely exclusively on the visual channel to share information. It has been suggested (Schmidt et al, 2008) that this lack of multisensory capability helps to explain why interacting with machines often proves far more difficult than person-to-person communication. This being the case, making use of the haptic and/or acoustic channels as well as the visual one could be very effective and the associated possibilities thus merit more serious investigation.

A multisensory approach in the rail sector

Operators in the rail sector, which is characterised by high levels of automation and centralisation, have to be able to rely on information presented to them with no possibility of direct independent verification. A control operator in a control centre, for example, cannot check the current status of operation visually by simply looking at the situation on the ground: the information displayed on the MMI system is the only source available. Also track maintenance teams have to be able to trust that speed restriction sections are properly displayed on their computer console.

A method for generating a reliable display as described in Forstreuter et al (1994) using visual, aural and haptic information is described in the following. The method uses the human capacity to integrate different sensory inputs cognitively (Seebode, 2010), that is to say to compare the content of information received via the visual channel with that received via the auditory and haptic channels. The object of the invention is to improve a method for presenting an image on a display. Currently, process data is converted to image data and then put on a monitor. Ensuring that the image data is correct and reliable involves very considerable technical effort.
The development described in this paper makes it possible for the operator to detect inaccuracies in the information displayed without complex electronic systems. This involves passing the process data through a data processing step to generate a computed graphical representation as well as the image data. If the user then selects a display range on the screen, a check can be performed using a fault feedback function to determine whether there is any discrepancy between the image element displayed and the corresponding image element in the computed representation. An example is provided below to explain how this works in greater detail.

Figure 1 shows a railway signaling safety system (1) from which process data (P) is output via communication channels (2 and 3). This process data is supplied to data processing devices (4 and 7). An image showing multiple tracks, switches and signals is generated on a touchscreen (6) via a communication link (5), while a computed representation is created in (7) and stored in electronic form. This data processing device (7) is capable of providing both acoustic and haptic feedback (vibration). For example, if the user now touches a red signal on the touchscreen (6) that conforms to the corresponding image element in the computed representation, an acoustic signal and/or a haptic signal is generated. However, if there is a discrepancy with respect to the corresponding image element in the computed representation, no sound and/or vibration is generated.

Figure 1: Schematic diagram of the method for detecting erroneous monitor displays
As described, this method differs from others in that it uses no technical safety modules and instead entrusts the role of safety comparator as according to Forstreuter et al (1994) to the human participant, who is actively involved in the decision chain.

It is not indicated whether the sensory mode – haptic, acoustic or both – used to report the conformity of the image elements, makes any difference. It is thus advisable, prior to employing the method in practice, to investigate whether there are any discernible differences in terms of detection capability on the part of the user and his/her expected failure rates when using the combinations

- visual and haptic information (VH),
- visual and acoustic information (VA) and
- visual, haptic and acoustic information (VHA).

It was found in Pomper et al (2014) that a multisensory approach can have an effect on the response. Users were found to respond significantly better to the multisensory VHA signal than to a single-channel signal. This is confirmed by Ng and Chan (2014), which also shows that differences occur in the response to single-channel signals.

Given the many influencing factors involved, there is consequently good reason to examine the hypothesis that there exists a significant difference between the different feedback modes.

Experience suggests that, in the case of frequently recurring signals, haptic feedback tends to be less disruptive than a sound. If the user is holding the tablet PC, for example, it will not generate any noises capable of causing disturbance in the surroundings in a way that acoustic feedback very easily can (Lumbsden, 2008). A headset could achieve the same effect, but it is doubtful that a user would want to wear a headset permanently. Sound signals, moreover, ideally need to be emitted from the location concerned, as actions can be performed faster and more accurately if the spatial position coincides with the trigger point (Hommel, 2014). This is not the case when using a headset.

Two further considerations also suggest that haptic feedback might be more suitable to help the operator reach a correct decision. Vibration can be perceived clearly even in the presence of conversations or noise, up to a certain threshold, so this option increases the potential area of application (although it must be remembered that the busier the environment becomes, the more easily the user can be distracted). Studies (Bargedick et al, 2011) have found, furthermore, that the proportion of information that is retained in the memory is up to 40% higher when information is imparted haptically, rather than acoustically or visually.
Description of the studies

The different multisensory signal combinations were researched in two studies. The first study used a game to test the suitability of the different feedback modes and attempts to determine a trend. The second study brought the investigations closer to the railway operation sphere by testing the feedback modes with a simple operator workstation simulation.

Test 1 “Bees and Flowers”

This test was conducted in two different types of environment. The first environment was a busy foyer of a Siemens building at the Braunschweig site, where 46 subjects ranging in age from 17 to 60 (median 33.5) took part in the study. All the subjects were randomly selected employees of Siemens AG at the Braunschweig site. The second environment used was a quiet office environment, where 21 subjects ranging in age from 9 to 52 (median 23) took part in the study. Most of the subjects were employees or students of Braunschweig University of Technology.

A between-subjects design was selected for both environments. Each subject in the foyer environment worked with one of the three different modes VH, VA and VHA. An additional mode, double visual (VV), was investigated as well in the office environment. The modes were changed after each subject, such that the three/four modes VH, VA, VHA and VV were covered iteratively. The three/four feedback modes thus represented the independent variable.

The subjects were tasked with detecting erroneous bee and flower symbols. A bee was considered to be correct if touching the bee symbol triggered haptic (device vibrates), acoustic (device emits buzzing sound), combined haptic and acoustic (device vibrates and emits buzzing sound) or visual (additional symbol displayed) feedback, depending on the mode being tested. A flower was considered to be correct if no feedback was triggered. The artificial errors the subjects were tasked with detecting were thus a bee with no feedback and a flower with feedback. Subjects had to complete 30 levels, with twelve symbols being displayed at each level. The probability of each symbol being a bee or a flower was random (p = 0.5). The probability of one or more symbols at a level being erroneous was also set to p = 0.5. Subjects were instructed to press a confirmation button if they believed that all the symbols displayed at a level were correct. They were to press an error reporting button as soon as they detected an erroneous symbol without checking the remaining symbols. A preliminary test found that subjects often failed to comply with this rule, but checked further elements before reporting a failure. The failure rates were therefore recorded for analysis according to both the strict immediate reporting rule and a more flexible rule that recognised errors as having been detected provided they were reported at some point.

The relative frequency of the two possible failure types – failing to detect an erroneous signal or falsely reporting a correct one – served as the dependent
variable. The tests were conducted using the Samsung Galaxy Tab GT-P1000
tablet PC. The test process began with a short tutorial in which the subject was
told about the task and given the opportunity to sample touching or hearing
correct or erroneous symbols. Subjects then began to work through the test
independently, which generally took them about three minutes. Having
completed the practical element of the test, subjects were asked to complete a
questionnaire recording demographic data and to give a subjective assessment of
the test results.

Test 2 “Operator Workstation Simulation”
This test was conducted with twelve subjects ranging in age from 22 to 45
(median 24). Most of the subjects were employees of Siemens AG at the
Braunschweig site. The test employed a within-subjects design in which each
subject worked through the three applications one after the other. The task in this
case was to detect erroneous red signals in a prototype operator workstation
simulation. Every three to six seconds, a random signal would change colour.
The probability of this becoming erroneous was p = 0.2. It was necessary to
test all signals continuously in order to detect errors. The sequence in which
they were presented was based on drawing lots, with each of the six possible
sequences being completed twice. If the signal was correct, acoustic (VA), haptic
(VH) or combined haptic and acoustic (VHA) feedback was triggered, dependi-
g on the application. An absence of this feedback indicated an artificially genera-
ted error, which subjects were instructed to report using the error reporting button.
The three feedback modes thus represented the independent variable. The
relative frequency of the two possible failure types – failing to detect an
erroneous signal or falsely reporting a correct one – served as the dependent
variable. The tests were conducted using the Samsung Galaxy Tab GT-P1000
tablet PC. Each subject received a brief verbal introduction and description of the
task and was then allowed 30 seconds to try out the application and ask any
questions. The subject was then informed that data recording was starting.
Subjects worked through the three applications in succession spending four
minutes on each. Having finished the practical element of the test, subjects were
asked to complete a questionnaire recording demographic data and to give a
subjective assessment of the different feedback modes.

Results

Test 1 – “Bees and Flowers”
The results obtained in the foyer test environment are described below. The rate
of erroneous signal detection failures (erroneous signal not detected) did not
follow a normal distribution, so the nonparametric Kruskal-Wallis test was used.
There was no significant difference between the various feedback modes, $\chi^2(2) =
1.61$, $p = .45$. The rate of erroneous signal detection failures showed almost no
variation across the "haptic" (median = .063), "acoustic" (median = .029) and
"haptic and acoustic" (median = .071) conditions. The average rate of erroneous
signal detection failures under the strict rule was slightly lower for haptic
feedback (mean value = .375) than for combined feedback (mean value = .459)
and acoustic feedback (mean value = .468). The variance analysis for independent samples was not significant, F(2,43) = .67, p = .52. The rate of false alarms did not follow a normal distribution either, so here too the nonparametric Kruskal-Wallis test was used. There was no significant difference between the various feedback modes, χ²(2) = .155, p = .93. The median for all feedback modes was .000.

The rate of erroneous signal detection failures in the office environment showed almost no variation across the "haptic" (median = .028), "acoustic" (median = .000), "acoustic and haptic" (median = .000) and "visual" (median = .063) conditions. The average rate of erroneous signal detection failures under the strict rule was slightly lower in the case of combined feedback (median = .083) than for haptic (median = .254), visual (median = .333) and acoustic (median = .400) feedback. The median for false alarms across all feedback modes was .000. The study thus did not strengthen the hypothesis that the different feedback modes produce different error rates.

**Test 2 – “Operator Workstation Simulation”**

The rate of erroneous signal detection failures (erroneous signal not detected) did not follow a normal distribution. The necessary conditions for the use of normal distribution methods were thus not met and the nonparametric Friedman test was consequently used. There was no significant difference between the various feedback modes, χ²(2) = .00, p = 1.0. The medians amounted to .027 for the acoustic mode, .033 for the haptic mode and .034 for the combined acoustic and haptic mode. The differences are thus very small. The rate of false alarms did not follow a normal distribution either. The Friedman test produced no significant result, χ²(2) = .56, p = .76. The rate of false alarms showed almost no variation across the "haptic" (median = .003), "acoustic" (median = .004) and "haptic and acoustic" (median = .034) conditions.

Removing the effects of outliers changed the results only slightly. The mean number of failures is lower for the haptic feedback mode (median = .028) than for the acoustic (median = .049) and combined modes (median = .048). The result of the statistical test is consequently also supplied to inform the assessment as to whether there is a difference between the feedback modes. The variance analysis for dependent samples was not significant, F(2) = 1.288, p = .3. Thus this study did not strengthen the hypothesis that the different feedback modes produce different error rates.

Assessments of the effort required to remain alert during the task varied from subject to subject. The average choice on the five-point scale from "High" to "Minimal" was the middle. The subjects subsequently rated "acoustic", "haptic" and "haptic and acoustic" feedback on a five-point scale from "Very poor" to "Very good". Their ratings of the feedback modes are presented in Figure 2.
The number of subjects awarding each rating was extracted. The results of this evaluation showed a wide variation. None of the feedback modes was rated "Very poor". Acoustic feedback, as the diagram shows, received the poorest rating. If the possible responses are scored 1 for "Very poor" through to 5 for "Very good", the medians for acoustic, haptic and the acoustic/haptic combination equate to 3, 4 and 4.5 respectively, with interquartile ranges of 2.0, 0.75 and 1.75 respectively. Acoustic feedback was thus judged to be moderate, haptic feedback good and combined acoustic and haptic feedback good to very good. Evaluation of the "acoustic feedback", "haptic feedback", and "acoustic and haptic feedback" modes did not follow a normal distribution, so the nonparametric Friedman test was used. A significant difference was determined between the various feedback modes, $\chi^2(2) = 8.77$, $p = .01$. This study thus strengthens the hypothesis that the different feedback modes are evaluated differently.

**Conclusion**

The studies described can provide an initial indication as to whether and under what circumstances different feedback modes are effective. A brief discussion of the results is presented below along with a review of possible sources of error and unintended influencing factors. The authors acknowledge that other factors not addressed here may also have affected the results.

The assessment of the feedback modes based on the "Bees and Flowers" application showed that the results appear qualitatively equivalent whichever sensory channel is utilised. Neither the foyer test results nor the office
environment test results support any conclusion as to the superiority of one feedback mode over another. Test supervisors observed that many subjects completing the tests adopted a particular systematic approach to the task after just a few levels. Some subjects checked all the bees first and then all the flowers (or vice versa), possibly in order to be able to recognise more easily whether or not the presence or absence of feedback indicated an error. Other subjects always checked the elements in the same order. Having the twelve elements arranged in the same way every time could have led subjects to adopt a fast pace such that some errors were only noticed after the subject had moved on to the next symbol. Delayed error detection for this reason could explain the high failure rate under strict application of the rules. This type of failure may also indicate that motor control and cognitive processing were proceeding asynchronously during error detection.

The small difference between visual and other feedback modes suggests, moreover, that the visual channel supports thinking no more effectively than other channels. Given the complexity of the task, it can be assumed that the effect of the multisensory approach on decision-making processes in the brain was small. A subsequent study could investigate whether the haptic and/or acoustic modes might perform better than the visual mode when the visual load is high.

The low failure rates and positive assessment of the feedback modes from the tests using the operator workstation simulation indicate that the application can be considered fit for use and that the human sensory system is compatible with the haptic and acoustic display developed. Subjects reported following the test that they had difficulties selecting the signals because the tablet PC display was too small. Some subjects also noted that the sound component of the combined haptic and acoustic feedback mode was very slightly delayed. Both of these factors could have led to operating errors, and using a larger tablet PC with improved sensitivity that emits sounds with no delay, for example, may therefore reduce the number of errors further. As the user preference should be taken into account, the most effective feedback mode is determined by combining the evaluations of the failure frequencies and the preferences. The low number of failures means that all feedback modes can be considered for use. Users preferred haptic and combined feedback modes to the acoustic mode.

Summary

The studies have shown that the idea to use a variation of feedback modes to reveal mistakes on a display works. Different combinations of haptic, acoustic, visual and combined haptic and acoustic feedback were researched.

While the studies have shown that potential users tend to prefer a haptic or combined haptic and acoustic signal over an acoustic signal alone, they were not able to ascertain any significant difference in the effectiveness of the various feedback modes.
References


Dose, D. 2014, *Evaluation der Eignung menschlicher Sensorik für eine beispielhaft entwickelte haptische und akustische sichere Anzeige im Eisenbahnkontext* [Evaluation of the suitability of the human sensory system for a model haptic and acoustic display developed for the railway operation context], (Chemnitz)


Hampel, T. 2011, *Untersuchungen und Gestaltungshinweise für adaptive multifunktionale Stellteile mit aktiver haptischer Rückmeldung* [Investigations and designs for adaptive multifunctional actuating components with active haptic feedback], (Stuttgart University Library)


