NHK STRL Bulletin
BROADCAST TECHNOLOGY
Winter 2010
NHK Science & Technology Research Laboratories
ISSN 1345-4099
no. 39

CONTENTS

On the 80th anniversary of STRL

Keiichi Kubota, *Director General*

This year, the NHK Science and Technology Research Laboratories celebrates its 80th anniversary. STRL was established in 1930, only five years after the beginning of radio services in Japan. The initial objective of the organization was to develop the infrastructure needed for television broadcasting of the Tokyo Olympic Games planned to be held in 1940. Although the Olympics was canceled because of the war, the enthusiasm that researchers took to their work in those days has become woven into the DNA of today's STRL researchers. Although we can pause to reminisce about the history of STRL, we should also envision what the coming 20 years will hold for television and look towards our 100th anniversary.

All employees of NHK are now doing their best to fulfill the Corporate Plan[1] for 2009-2011. STRL researchers are working hard to embody the concept of "Wherever you are, NHK", which emphasizes services for "3-Screens", TVs, PCs and cellular phones, as well as research on fusing broadcasting and communications media and human friendly technologies.

Concerning the "3-Screens" concept, besides developing various technologies and applications, this year, we will demonstrate the services and home terminals we are planning for the broadcasting - communication era. Furthermore, for Super Hi-Vision, the flagship of our research projects, we will begin the switchover from the dual-green format to the full 33Mega pixel resolution system. We started a joint project with the BBC aiming at public viewing of Super Hi-Vision at the 2012 London Olympic Games. Of course, such big projects cannot be realized by STRL alone. We hope for support and cooperation from various groups.

At this year's Open House, we will review STRL's history and make our ten-year and twenty year plans public along with our latest research results. Your comments and suggestions are as always greatly appreciated. Thank you.

1: http://www.nhk.or.jp/pr/koho-e.htm

NHK STRL

Feature

Outline of Barrier-Free Broadcasting Technologies for Visually-Impaired People

Recent developments in information and communication technology (ICT) are enriching our lives with the convenience of a huge variety of information services. At the same time, however, this has produced an information gap in that disabled and elderly people are unable to use and apply such services to the fullest extent. That is why we at NHK STRL are actively pursuing research and development into technologies aimed at "information barrier-free" services, to ensure that everyone, including elderly or disabled people, can enjoy a rich variety of advanced broadcasting services. In this article, we take up the challenge of providing information barrier-free technologies in broadcasting for visually-impaired people, and also comment on current research into touch presentation technologies and audio presentation technologies that NHK is working on.

1. Introduction

Recent rapid developments in ICT are causing huge changes in our lives. We can now access a large variety of information wherever and whenever we want, and we can now expect lifestyles that are richer and highly convenient. But at the same time, an information gap is being created in that disabled or elderly people find it difficult to access content that is provided by broadcasts or over the Internet.

According to the results of a survey into the status of physically disabled children and adults conducted by the Ministry of Health, Labour and Welfare in 2006, there are 310,000 people with visual impairments and 276,000 people with hearing impairments in Japan. It is also said that we will become a super-graying society by the year 2015, with one in four people being 65 or over, so it is thought that the number of people with visual and hearing disabilities will increase further in the future. Against that background, there is a call for an "information barrier-free" broadcasting service that will enable anyone to enjoy enriched and advanced broadcasting services, even disabled or elderly people.

As broadcasting becomes digital, it is becoming possible to transmit not only programs containing high-quality video images and sound, but also a large variety of data. At NHK, we are working towards expanding our closed-captioned broadcasting services for hearing-impaired people and commentated broadcasting services for visually-impaired people.

In 2007, the number of hours of closed-captioned broadcasting by NHK was 3,913 hours, 15 minutes (44.6%) on the General TV channel and 2,339 hours, 36 minutes (29.9%) on the Educational TV channel. The proportion of closed-captioned broadcast hours in the broadcast hours that could be tagged with closed-captioning reached 100% in 2006 on the General TV channel. The closed-captioning of live broadcast programs is progressing, by closed-captioned broadcasting of news programs in which speech recognition is used or by a respeak method that recognizes the speech by a separate speaker called a "closed-captioning newscaster" re-phrases the content, and closed-captioned broadcasting is expected to expand further in the future.

On the other hand, although the number of hours of commentated broadcasts, which is a broadcasting service for visually-impaired people, has increased each year, by 2007 it had only reached 328 hours, 29 minutes (3.7%) on the General TV channel and 685 hours, 23 minutes (8.7%) on the Educational TV channel. In addition, data broadcasts and electronic program guides (EPGs), which are new services in digital broadcasting, require the user to use a graphical user interface (GUI) to operate a remote control interactively while viewing a screen, so these are broadcasting services and receiver functions that visually-impaired people are virtually unable to use.

Thus the process of providing information barrier-free services for visually-impaired people has not progressed as far as that of information barrier-free services for hearing-impaired people. In this article we take up the challenge of providing information barrier-free technologies for visually-impaired people, and also comment on information barrier-free broadcast technologies for visually-impaired people that uses tactile and audio presentation technologies which NHK is working on.

Table 1: Numbers of centripetal fibers and amounts of information for each of the five senses

Sense	Number of centripetal fibers	Physiological information amount	Conscious perception information amount
Sight	$2x10^6$	10^7 bps*	40 bps
Hearing	$2x10^4$	10^5 bps	30 bps
Touch	10^6	10^6 bps	5 bps
Taste	10^3	10^3 bps	1 bps (unclear)
Smell	10^5	10^5 bps	1 bps (unclear)

*bits per second

2. The challenge of barrier-free technologies designed for visually-impaired people

TV content consists of visual information and aural information, so that the lack of either type of information will make it impossible to fully comprehend details. If there is a disability in the sensory receptors (or cerebral cortex region) corresponding to either type of information, so that one type of information cannot be received sufficiently, it is necessary to convert the sensory information affected by the disability into sensory information that can be received. If there are disabilities in both sight and hearing, the information must be converted and presented as another type of sensory information (such as tactile information). In other words, it is necessary to convert the sensory information as follows:

Visual disability: visual information → aural or tactile information

Hearing disability: aural information → visual or tactile information

Visual and hearing disabilities: visual or aural information → tactile information

Estimates of the amount of information processed per unit time by each of the five senses are given in Table 1. The physiological information amount is the maximum

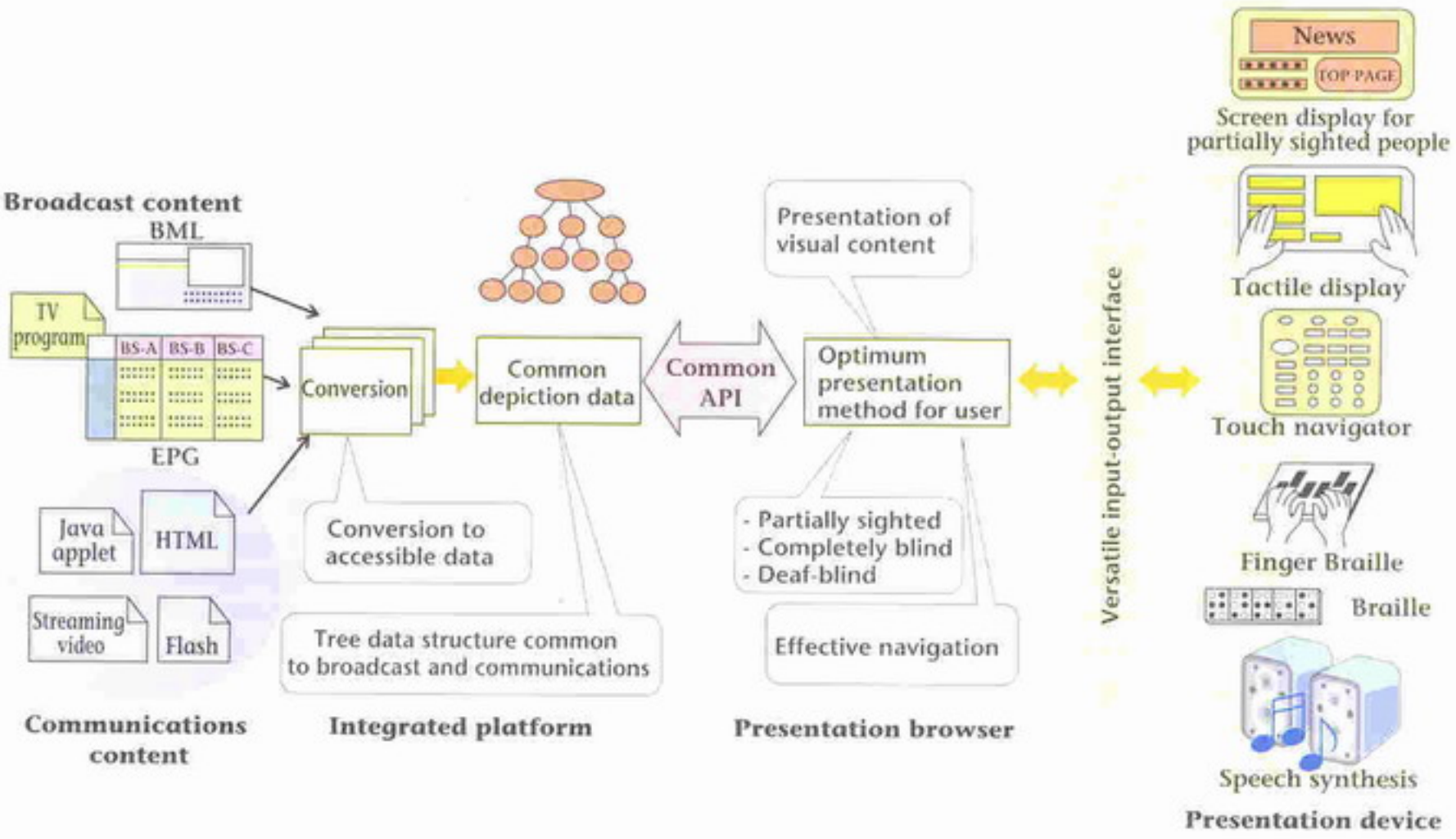

Figure 1: Overall flow of information barrier-free reception and presentation system

Table 2: Information presentation methods in information barrier-free reception and presentation system

Information Presentation Method	Main Target	Presented Information	Characteristics
Audio reading	Completely blind people (excluding deaf-blind people)	Verbal information	- Special training unnecessary - High-speed reading possible - Can't be used when there's no text information
Braille	Deaf-blind people	Verbal information	- Fast reading possible, depending on skill - People who can read Braille are approximately 10% of visually-impaired people
Six-finger Braille	Deaf-blind people, Braille beginners	Verbal information	- Fast reading possible, depending on skill - Easy to learn, effective for partially sighted people and Braille beginners
Touch pattern	Completely blind people	Verbal information, two-dimensional patterns	- Conversion to verbal information enables presentation of difficult information - Can be used as input-output interface - Effective for partially sighted people
Enlargement, color conversion	Weakly sighted people	Verbal information, two-dimensional patterns	- Enables display to suit type and degree of disability

amount of information, calculated from the numbers of centripetal fibers (neural fibers from the sensory receptors to the cerebral cortex) and the transmission capabilities of each neural fiber. The conscious perception information amount is the maximum amount of information for each sense that the user is conscious of. If the amount of visual and aural information obtained from TV content is close to the physiological information amount, there will be some leeway in the amount of information when aural information is converted into visual information or tactile information. However, when visual information that has a large amount of information is converted into aural information or tactile information, some form of information compression will be necessary. Thus the conversion of sensory information for visually-impaired people is expected to encounter more difficulties than the conversion of sensory information for hearing-impaired people.

When visual information is converted into aural information, it is usual to convert the visual information into verbal information, and then into audio information. The conversion from verbal information to audio information can be implemented by speech synthesis technology. However, to convert visual information directly into verbal information would require technology that recognizes (verbalizes) visual factors such as objects and their placements and behaviors within the video images, and many technological breakthroughs will be necessary before this could reach a practicable level. The scripts of commentated audio broadcasts that are currently being broadcast are written by scriptwriters having specialist skills. In order to automatically produce content for commentated audio broadcasts at the same level as these scripts, a higher level of intelligent graphics recognition technology will be necessary. We need some means of extracting important information that should be conveyed from within a video image, as verbal information, then reducing that amount of information to a level that the user is capable of receiving as audio information.

3. Information barrier-free broadcast technologies for visually-impaired people

We are working on developing information barrier-free technologies aimed at data broadcasting that enable the users themselves to control the speed at which information is received, by interactive operations. We have also set up the development of a service that will read out the superimposed captions of newsflashes about earthquakes and tsunamis, where there is a lot of verbal information in the data broadcasting content and a strong demand for turning it into audio information. Outlines of each of these technologies are given below.

3.1 Information barrier-free reception and presentation system

In data broadcasting, which is one of the digital broadcasting services, it is necessary to convey content that is expressed visually in such a manner that visually-impaired people can receive it too. In this case, we structure data broadcast content that has been described in Broadcast Markup Language (BML), add semantic information to it, and also convert it into a description format that can be used in common with communications content described in Hyper Text Markup Language (HTML). Using a common Application Program Interface (API), we have developed an information barrier-free reception and presentation system that can provide information from a wide variety of presentation devices (configurations) which are suited to different types and degrees of disabilities (see Figure 1 and Table 2).

Using BML content that had actually been broadcast by NHK, we confirmed the performance of the method of converting data into common depiction data, and as a result were able to apply sufficient structure and meaning to substantially all of the pages. In addition, from experiments to evaluate the versatility of the conversion rules necessary for the conversion, it was clear that the same conversion rules could be utilized continuously, provided the page design was not modified. Furthermore, the results of evaluations of accessibility to data broadcasting and EPGs, using presentation devices that have been adapted for each category of partially sighted, completely blind, and deaf-blind people, showed that information could be acquired easily and was easy to understand.

3.2 Touch presentation technology

We have prototyped a tactile display that is provided with an optical touch panel, for conveying visual content such as figures, graphs, and GUI screens in an effective manner, as an information presentation device for the information barrier-free reception and presentation

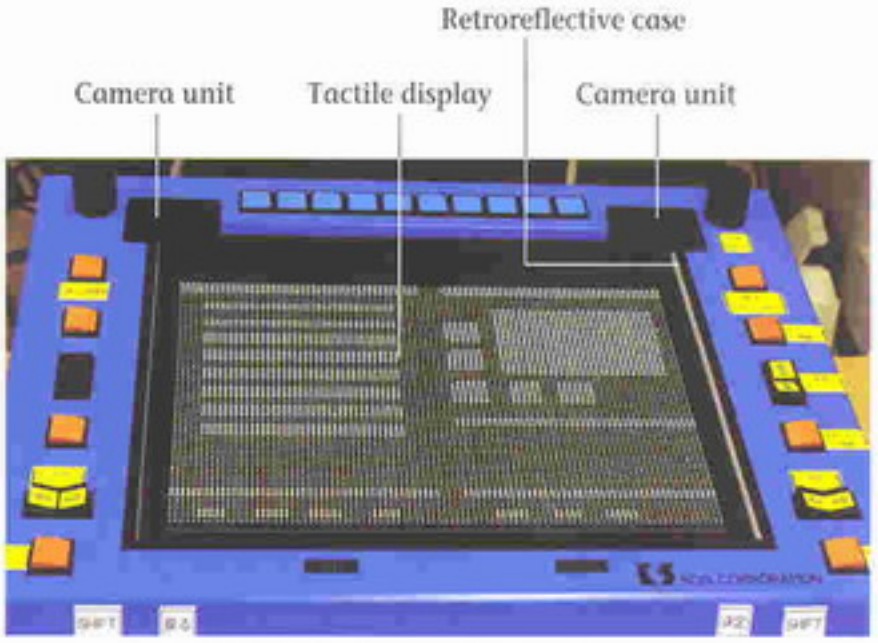

Figure 2: Tactile display with optical touch panel

system (Figure 2). The precision of positions when objects presented by the raising and lowering of pins are felt with the fingers is thought to depend on the mechanical precision of sensors and human movement errors, but the influence of the mechanical precision of the sensors was small. In this case, as the result of evaluating whether or not the size and placement conditions of objects can be selected accurately, it was clear that objects could be selected with an accuracy of at least 90% at sizes greater than the 2x3 dots (where 1 dot = 2.4 mm) that is equivalent to a single Braille character, provided the spacing between adjacent dots was at least 1 dot. We asked completely blind people to use this tactile display to search through structured content, and investigated search times and ease of operation. As a result, we found that although search times were somewhat longer at the start, in comparison with an interface with conventional key operations, the search time shortened rapidly due to the learning effect, and the system was evaluated as making it easy to comprehend the structures of data, even of tables and EPGs. In addition, basic research is under way into clarifying tactile perceptions of presented conditions and shape, to give the user a tactile sensation that is close to that given by the real object, as research into conveying a three-dimensional shape by touch. We are currently developing an experimental device that enables a user to hold a virtual object by six fingers of both hands, to enable a virtual impression of touching it.

3.3 Audio broadcasting services for visually-impaired people

We are conducting research into an audio broadcasting service that automatically reads out the captions of newsflashes displayed on the TV screen, with high-quality synthesized speech. We are developing a system that uses

the mechanism of data broadcasting to read out data that is generated by broadcast stations on the basis of earthquake and tsunami information issued by the Japan Meteorological Agency. We prototyped two different systems, one using a method of broadcasting a synthesized reading track on a supplementary sound channel of digital broadcasting and the other using a method of broadcasting information as a data file in data broadcasting. For the data file method, we performed behavior validation experiments using an actual operational configuration, the same transmitting equipment as that used by broadcast stations in practice, and digital broadcast receivers that major Japanese manufacturers have put on sale, and confirmed that such a setup conforms to that of commercial receivers.

4. Conclusion

The spread of broadband for Internet access has added content that provides not just text-based information, but also video images and audio information. In addition, mobile phones are enabling communications that are not just audio information, but also textual information by mail and video image information by TV phones. It is not yet clear whether or not it will be sufficient to provide an environment that allows the selection of sensory information suited to the type and degree of disability, to enable access to information and facilitate communication, but there has been steady progress when compared with the situation a few years ago.

In the broadcasting service field, some people say "an environment is in place to enable people with disabilities to use services that are appropriate for their disabilities in order to access information, such as TV and closed-caption broadcasting for hearing-impaired people and radio broadcasting for visually-impaired people". However, the results of the survey into the status of physically disabled children and adults conducted by Japan's Ministry of Health, Labour and Welfare in 2006 showed that the most common medium used by visually-impaired people for obtaining information is the TV. A similar result was obtained by a survey done by the Japan Federation of the Blind in 2004. From the feeling that many users, including able-bodied people, consider that they would like enjoy content together, with information in common, we can infer that they "listen" to the TV, which is watched by 90% of Japanese people out of all the various different media and has the highest contact ratio. In other words, when people actively access information they want to know about, or when they communicate in a one-to-one manner, they select the medium with the optimum sensory information. But with TV, it seems that most people would prefer the communal factors of "enjoyable, informative" over the reception of content with the optimum sensory information.

To enable us to respond to these needs of disabled people, it is necessary to construct a broadcast system that can provide a wide variety of sensory information with respect to the same content. This encourages us to develop sensory information conversion technologies and information presentation technologies that are suitable for various different disabilities, with the goal of an information barrier-free broadcasting service.

(Toshiya Morita and Nobuyuki Yagi)