

## Review Article

# Image and Video Processing for Visually Handicapped People

Thierry Pun,<sup>1</sup> Patrick Roth,<sup>1</sup> Guido Bologna,<sup>2</sup> Konstantinos Moustakas,<sup>3</sup> and Dimitrios Tzovaras<sup>3</sup>

<sup>1</sup> Computer Science Department, University of Geneva, Battelle Campus, 7 Route de Drize, 1227 Carouge (Geneva), Switzerland

<sup>2</sup> Computer Science Department, University of Applied Studies (HES-SO), 4 Rue de la Prairie, 1202 Geneva, Switzerland

<sup>3</sup> Center for Research and Technology Hellas (ITI/CERTH), Informatics and Telematics Institute, 1st Km Thermi-Panorama Road, P.O. Box 361, 57001 Thermi-Thessaloniki, Greece

Received 30 November 2007; Accepted 31 December 2007

Recommended by Alice Caplier

This paper reviews the state of the art in the field of assistive devices for sight-handicapped people. It concentrates in particular on systems that use image and video processing for converting visual data into an alternate rendering modality that will be appropriate for a blind user. Such alternate modalities can be auditory, haptic, or a combination of both. There is thus the need for modality conversion, from the visual modality to another one; this is where image and video processing plays a crucial role. The possible alternate sensory channels are examined with the purpose of using them to present visual information to totally blind persons. Aids that are either already existing or still under development are then presented, where a distinction is made according to the final output channel. Haptic encoding is the most often used by means of either tactile or combined tactile/kinesthetic encoding of the visual data. Auditory encoding may lead to low-cost devices, but there is need to handle high information loss incurred when transforming visual data to auditory one. Despite a higher technical complexity, audio/haptic encoding has the advantage of making use of all available user's sensory channels.

Copyright © 2007 Thierry Pun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. INTRODUCTION: VISUAL HANDICAP AND ASSISTIVE DEVICES

Visual impairment can be quantified in terms of the remaining visual acuity and visual field. Visual acuity is expressed as a fraction of full acuity; for instance, a visual acuity of 1/10 means that a sight handicapped person has to be at 1 meter to properly see an object seen at 10 meters by a normally sighted individual. The visual field is expressed in degrees; a normally sighted person is considered to have a visual field of about 60 degrees.

A distinction is made between low vision, legal blindness, and total blindness. According to the 10th Revision of the World Health Organization (WHO) International Statistical Classification of Diseases, Injuries and Causes of Death, low vision is defined as visual acuity of less than 6/18, but equal to or better than 3/60, or corresponding visual field loss to less than 20 degrees, in the better eye with best possible correction. The definition of legal blindness varies according to countries; it is usually stated as a visual acuity of less than 3/60, or corresponding visual field loss to less than 10 degrees, in the better eye with best possible correction. Total blindness means no remaining visual perception at all, and can be congenital or late. The fact that there is no visual per-

ception does not necessarily imply that the entire visual pathway (from the eye and retina to the cortex) is ineffective; in fact this is most often not the case. Following a 2002 survey, the WHO estimated that there were 161 million (about 2.6% of the world population) visually impaired people in the world, of whom 124 million (about 2%) had low vision and 37 million (about 0.6%) were blind [1]. Still according to WHO, more than 90% of the world's visually impaired live in developing countries, and more than 82% of all people who are blind are 50 years of age and older although they represent only 19% of the world's population.

According to these figures, a vast number of persons are therefore affected by some form of visual handicap. Various devices exist to assist their needs, to accomplish daily routine tasks at home, at work, or when traveling. Even very simple aids like the long cane, spelling watches, embossed documents, tactile and audio signposts, and so on can be tremendously helpful and have gained very wide acceptance. More sophisticated apparatus exist commercially or in laboratories, which very often perform some form of image/video processing to extract pertinent information from a visual signal. The need for image/video processing comes from the fact that the fundamental goal of these assistive aids is to complement or replace sight by another modality. The visual

information therefore needs to be simplified and transformed in order to allow its rendition through alternate sensory channels, usually auditory, haptic, or auditory-haptic.

As the statistics above show, the large majority of visually impaired people is not totally blind but suffers from impairments such as short sight which decreases visual acuity, glaucoma which usually affects peripheral vision, age-related macular degeneration which often leads to a loss of central vision. Due to the prevalence of these impairments and henceforth the need for mass-produced aids, these low-vision (as opposed to blindness) aids are often of low technicality; examples are magnifiers, audio books, and spelling watches. In other words, the need for mass-market computerized devices with image processing capabilities is not strongly felt. There are some exceptions, such as screen readers coupled with zoom (possibly directly operating on JPEG and MPEG data) and OCR capabilities, possibly with added Braille and/or vocal output. A noted work was the development of the Low Vision Enhancement System (LVES), [2]) where significant efforts were put on portability, ergonomics, and real-time video processing. A head-mounted display with eye tracking was used; processing included spatial filtering, contrast enhancement, spatial remapping, and motion compensation. Other systems include ad hoc image/video processing that compensates for a particular type of low-vision impairment. Typical techniques used are zooming, contrast enhancement, or image mapping (e.g., [3–7]). These low-vision aids will not be described further in this article, which concentrates on aids for the totally blind. Note however that some of the devices for the totally blind also target partially sighted users.

One of the long-term goals of research in the domain of assistive aids for blind persons is to allow a totally sightless user to perceive the entire surrounding environment [8–10]. This not only requires to perform some form of scene interpretation, but also that the user is able to build a mental image of his/her environment. An important factor to take into consideration is then the time of appearance of blindness, from birth or later. According to [11], mental images are a specific form of internal representation, and their associated cognitive processes are similar to those involved in other forms of perception. The mental image is obtained according to an amodal perceptual process. The term “amodal” has been established following several studies made on congenitally blind people, which proved that a mental image is not uniquely based on visual perception [12]. In the case of a blind person, the mental image is usually obtained through the use of haptic and auditory perception. Kennedy [13] claimed that congenitally blind subjects could recognize and produce tactile graphic pictures including abstract properties such as movement. He also recognized that blind people are also able to understand and utilize perspective transformations, which is contested by [14]. According to Arditi’s point of view, congenitally blind users cannot access purely visual properties, which include the perspective transformations. Studies reported by [15] also revealed that congenitally blind people were able to generate and use mental images from elementary tactile pictures. However, they suffer from imagery limitations when tactile images increase in complexity. These

limitations are caused by their spatial perceptual deficit due to their blindness and the high attentional load associated with the processing of spatial data. Hatwell also assumed that haptic spatial perceptions from congenitally blind people are systematically less efficient than for the late blind persons. This comes from the visual-haptic cross-modal transfer that took place during the infancy of the late blind and which increased the spatial perceptual quality of such sensory system. This set of observations showed that early and late blind people are able to generate mental images, although the process is harder for early blind persons. Furthermore, associations of colors to objects will only be known at an abstract level for people having never experienced sight. In any case, the content of nonvisual pictures must be previously simplified, in order to minimize the cognitive process necessary for the recognition.

Another very important issue is the development of orientation, mobility, and navigation aiding tools for the visually impaired. The ability to navigate spaces independently, safely, and efficiently is a combined product of motor, sensory, and cognitive skills. Sighted people use the visual channel to gather most of the information required for this mental mapping. Lacking this information, people who are blind face great difficulties in exploring new spaces. Research on orientation, mobility, and navigation skills of people who are blind in known and unknown spaces, indicates that support for the acquisition of efficient spatial mapping and orientation skills should be supplied at two main levels: perceptual and conceptual [16, 17].

At the perceptual level, the deficiency in the visual channel should be compensated by information perceived via other senses. The haptic, audio, and smell channels become powerful information suppliers about unknown environments. Haptics is defined in the Webster dictionary as “of, or relating to, the sense of touch.” Fritz et al. [18] define haptics: “tactile refers to the sense of touch, while the broader haptics encompasses touch as well as kinaesthetic information, or a sense of position, motion, and force.” For blind individuals using the currently available orientation, mobility, and navigation aids, haptic information is commonly supplied by the white cane for low-resolution scanning of the immediate surroundings, by palms and fingers for fine recognition of object form, texture, and location, and by the feet regarding navigational surface information. The auditory channel supplies complementary information about events, the presence of others (or machines or animals), or estimates of distances within a space [19].

At the conceptual level, the focus is on supporting the development of appropriate strategies for an efficient mapping of the space and the generation of navigation paths. Research indicates that people use two main scanning strategies: route and map strategies. Route strategies are based in linear (and therefore sequential) recognition of spatial features, while map strategies, considered to be more efficient than the former, are holistic in nature. Research shows that people who are blind use mainly route strategies when recognizing and navigating new spaces, and as a result, they face great difficulties in integrating the linearly gathered information into a holistic map of the space.

The reminder of this article is organized as follows. Section 2 presents the main alternate sensory channels that are used to replace sight that is haptic, auditory, and their combination. Direct stimulation of the nervous system is also discussed. The following sections then review various assistive devices for totally blind users, classified according to the alternate modality used. This classification was preferred over one based on which image/video processing techniques are employed, as many systems use a variety of techniques. It was also preferred over a description based on the situation in which a device would be used, since various systems aim at a multipurpose functionality. Section 3 thus discusses systems relying on the haptic channel, historically the first to appear. Section 4 concerns the auditory channel, while Section 5 discusses the use of the combination of auditory and haptic modalities. Section 6 discusses these devices from a general viewpoint and concludes the article.

## 2. ALTERNATE SENSORY CHANNELS AND MODALITY REPLACEMENT FOR THE TOTALLY BLIND

Sight loss creates four types of limitations, regarding communication and interaction with others, mobility, manipulation of physical entities, and orientation in space (e.g., [20–22]). To compensate for total or nearly total visual loss, modality replacement is brought into play, which is the basic development concept in multimodal interfaces for the disabled. *Modality replacement can be defined as the use of information originating from various modalities to compensate for the missing input modality of the system or the users.*

The most common modality to replace sight is touch, more precisely the haptic modality composed of two complementary channels, tactile and kinesthetic [23]. The tactile channel concerns awareness of stimulation of the outer surface of the body, and the kinesthetic channel concerns awareness of limb position and movement. Haptic perception is sequential and provides the blind with two types of information that are of complementary nature: semantic “(what is it?)” and spatial “(where is it?)” [24]. Both types of information are at the end of the process combined together to form the mental image. Two strategies are applied when performing the exploration of a physical object using the hand, based on macro- and micromovements. Macromovements perform a global analysis, while micromovements consider details; assistive devices should therefore allow for these two types of explorations. The use of the haptic modality to replace the visual can be accomplished in two ways, that is, physically and virtually. In the physical interaction, the user interacts with real models that can be 3D map models or Braille code maps using the hands. In the virtual interaction, the user interacts with a 3D virtual environment using a haptic device that provides force/tactile feedback and makes the user feel like touching a real object. The physical haptic interaction is in general more efficient than the virtual due to the intuitive way of touching objects with the hands, instead of using an external device for interaction. However, virtual haptic interaction is more flexible and many 3D virtual environment and virtual objects can be rapidly designed, while with proper training it is reported that the user can easily

manipulate a haptic device to navigate in 3D virtual environments [25].

The other main replacement modality is hearing. Whereas touch plays the key role in the perception of close objects, hearing is essential for distant environments. A sound is characterized by its nature and its spatial location [26]. Monaural hearing can be sufficient in a number of situations, although binaural hearing plays an important role in the perception of distance and orientation of sound sources. Assistive devices that use the hearing channel to convey information should thus not prevent normal hearing; they should only become active at the user’s request (unless an alert needs to be conveyed). The audio and haptic modalities can also be used jointly, as is the case with some of the assistive aids that are presented below.

Research aiming at directly stimulating the visual cortex, thus bypassing alternate sensory channels, has been active for decades. Intracortical microstimulation is performed by means of microelectrodes implanted in the visual cortex (e.g., [27–29]). When stimulated, these electrodes generate small visual percepts known as phosphenes which appear as light spots; simple patterns can then be generated. An alternative approach consists in the design of artificial retinas (e.g., [30, 31]). In addition to technical, medical, and ethical issues, these devices require that at least parts of the visual pathways are still operating: the optical nerve in case of artificial retina as well as the visual cortex. Direct cortical or retinal stimulation will not be discussed further, but it should be noted that such apparatus call for sophisticated, real-time image processing to simplify scenes in such a way that only the most meaningful elements remain.

Regarding the available aids for the visually impaired, they can be divided into passive, active, and virtual reality aids. Passive aids are providing the user with information before his/her arrival to the environment. For example, verbal description, tactile maps, strip maps, Braille maps, and physical models [17, 32, 33]. Active aids are providing the user with information while navigating, for example, Sonicguide [34], Kaspas [35], Talking Signs or embedded sensors in the environment [36], and Personal Guidance System, based on satellite communication [37]. The research results indicate a number of limitations in the use of passive and exclusive devices, for example, erroneous distance estimation, underestimation of spatial components and objects dimensions, low information density, or misunderstanding of symbolic codes used in the representations.

Virtual reality has been a popular paradigm in simulation-based training, game and entertainment industries [38]. It has also been used for rehabilitation and learning environments for people with disabilities (e.g., physical, mental, and learning disabilities) [39, 40]. Recent technological advances, particularly in haptic interface technology, enable blind individuals to expand their knowledge as a result of using artificially made reality through haptic and audio feedback. Research on the implementation of haptic technologies within virtual navigation environments has yielded reports on its potential for supporting rehabilitation training with sighted people [41, 42], as well as with people who are blind [43, 44]. Previous research on the use of haptic devices

by people who are blind relates to areas such as identification of objects' shape and textures [45], mathematics learning and graphs exploration [46, 47], use of audio and tactile feedback for exploring geographical maps [48], virtual traffic navigation [49], and spatial cognitive mapping [50–52].

### 3. HAPTIC ENCODING FOR VISION SUBSTITUTION

#### 3.1. Tactile encoding of scenes

As seen above, two fairly different haptic modalities can be used: tactile and kinesthetic. Tactile devices are likely the most widely used to convey graphic information. Historically, the first proposed system dates back to 1881 when Grin [53] proposed the Anoculoscope. This system should have projected an image on an 8 by 8 array of selenium cells which, depending on the amount of impinging light should have controlled electromechanical pin-like actuators. This system was however never actually realized, “for lack of funding” as the inventor stated.

Coming to more recent work, some guidelines should be followed in order for an image to be transformed in a form suitable for tactile rendition. The tactile image should be as simple as possible; details make tactile exploration very difficult. Attention should be paid to the final size of objects; some resizing might be necessary. Crossings of contours should be avoided, by separating overlapping objects; contours should be closed. Text, if present, should be removed or translated into Braille. As image processing practitioners know, performing such image simplification is no mean feat and various solutions have been proposed (e.g., [54–56]). They have in common a chain of processing that includes denoising, segmentation, and contours extraction. Contours are closed to eliminate gaps, and short contours are removed, resulting in a binary simplified image. In some cases, regions enclosed by closed contours have been filled in by textures.

A critical issue is then how to render these images in tactile form. Two families of supports coexist, allowing for either static or dynamic rendering. Static images are in general produced by means of specific printers that heat up paper on which a special toner has been deposited; under the influence of the heat, this toner swells and therefore gives a raised image. Such static raised images are routinely used in many places; often however these images are prepared by hand and little image processing is involved.

Supports that permit dynamic display of images can be mechanic-tactile with raising pins, vibrotactile, electrotactile where small currents are felt in particular locations, and so on, (see [57] for a comprehensive review). The earliest system using a head-mounted camera and dynamic display was the Electrophthalm from Starkiewicz and Kuliszewski, 1963, later improved to allow for 300 vibrating pins [58]. Around the same time was developed the TVSS—Tactile Vision substitution System, with a 1024 array of vibrating pins located on the abdomen of the user (e.g., [59, 60]). A noted portable device using a small dynamic display of 24 by 6 vibrating pins is the Optacon, first marketed in 1970 by Telesensory Systems Inc. (Mountain View, Calif, USA) [61], and used until recently (e.g., [62, 63]). The user could pass a small camera

over text or images, and corresponding pins would vibrate under a finger. In terms of image processing, in such systems using dynamic display the image transformation was based on simple thresholding of grey-level images, where the threshold could be varied by the user.

Purely tactile rendition of scenes using dynamic displays suffers from several drawbacks. First, the information transfer capacity of the tactile channel is inherently limited; not more than a few hundreds of actuators can be used. Secondly, such displays are technologically difficult to realize and costly; they are also difficult to use for extended periods of time. Finally, apart from reading devices such as the Optacon, real-time image/video scene simplification is needed which is difficult to achieve with real scenes. The tactile channel is therefore often complemented with the auditory channel, as described in Section 5.

#### 3.2. Tactile/kinesthetic encoding of scenes

The basic principle there is to provide the user with force feedback and possibly additional tactile stimuli. Such approaches have been made popular due to the development of virtual reality force-feedback devices, such as the CyberGrasp [64], the PHANTOM family of devices [65], the Phantograph [66], gaming devices, or simply the Logitech WingMan force-feedback mouse [67]. Force feedback allows rendering a feeling of an object, of a surface. For instance, [68] investigated different methods for representing various forms of picture characteristics (boundary or shape, color, and texture) using haptic rendering techniques. A virtual fixation mechanism allows following contours as if one was guided by virtual rails. When the force-feedback pointer is close enough to the line, this mechanism pulls the end effector towards the line. Surface textures were also rendered by virtual bump mapping.

Colwell et al. [44] carried out a series of studies on virtual textures and 3D objects. They tested the accuracy of a haptic interface for displaying size and orientation of geometrical objects (cube, sphere). They also studied whether blind people could recognize simulated complex objects (i.e., sofa, armchair, and kitchen chair). Results from their experiments showed that participants might perceive the size of larger virtual objects more accurately than of smaller ones. Users also may not understand complex objects from purely haptic information. Therefore, additional information such as from the auditory channel has to be supplied before the blind user can explore the object. Other studies reported by [49] tested the recognition of geometrical objects (e.g., cylinders, cubes, and boxes) and mathematical surfaces, as well as navigation in a traffic environment. Results showed that blind users are able to recognize more easily realistic complex objects and environments rather than abstract ones.

In [69], a method has been proposed for the haptic perception of greyscale images using pseudo-3D representations of the image. In particular, the image is properly filtered so as to retain only its most important texture information. Next, the pseudo-3D representations are generated using the intensity of each area of the image. The user can then navigate

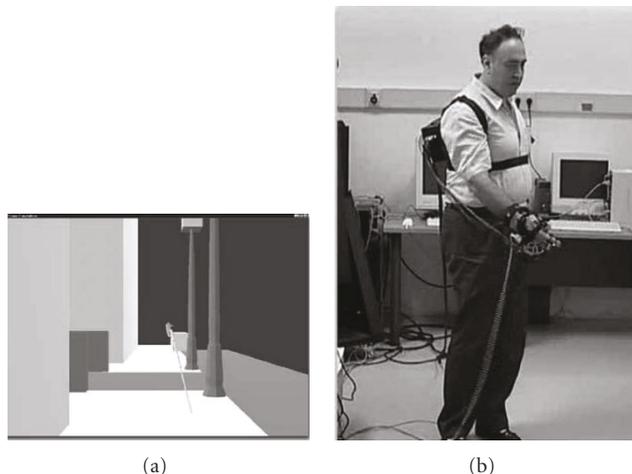


FIGURE 1: Cane simulation—outdoors test. (a) Virtual setup. (b) A user performing the test.

into the 3D terrain and access the encoded color and texture properties of the image.

Recently, Tzovaras et al. [25] developed a prototype for the design of haptic virtual environments for the training of the visually impaired. The developed highly interactive and extensible haptic VR training system allows visually impaired to study and interact with various virtual objects in specially designed virtual environments, while allowing designers to produce and customize these configurations. Based on the system prototype and the use of the Cyber-Grasp haptic device, a number of custom applications have been developed. The training scenarios included object recognition/manipulation and cane simulation (see Figure 1), used for performing realistic navigation tasks. The experimental studies concluded that the use of haptic virtual reality environments provides alternative means to the blind for harmlessly learning to navigate in specific virtual replicas of existing indoor or outdoor areas.

#### 4. AUDITORY ENCODING FOR VISION SUBSTITUTION

Fish [70] describes one of the first known works that used the auditory channel to convey visual information to a blind user. 2D pictures were coded by tone bursts representing dots corresponding to image data. Image processing was minimal. The vertical location of each dot was represented by the tone frequency, while the horizontal position was conveyed by the ratio of sound amplitude presented to each ear using a binaural headphone. At about the same time appeared a device, the “K Sonar-Cane,” that allowed navigation in unknown environments [71]. By combining a cane and a torch with ultrasounds, it was possible to perceive the environment by listening to a sound coding the distance to objects, and to some extent object textures via the returning echo. The sound image was always centered on the axis pointed at by the sonar. Scanning with that cane only produced a one dimensional response (as if using a regular cane with enhanced and variable range) that did not

take color into account. Some related developments used miniaturized sonars mounted on spectacles.

Later, Scadden [72] was reportedly the first to discuss the use of interface sonification to access data. Regarding diagrams, their nonvisual representation has been investigated by linking touch (using a graphical tablet) with auditory feedback. Kennel [73] presented diagrams (e.g., flowcharts) to blind people using multilevel audio feedback and a touch panel. Touching objects (e.g., diagram frames) and applying different pressures triggered feedback concerning information regarding the frame, and the interrelation between frames. Speech feedback was also employed to express the textual content of the frame. More recent works regarding diagrams presentation include for instance [74, 75]. Using speech output, Mikovec and Slavik [76] defined an object-oriented language for picture description. In this approach, an image was defined by a list of objects in the picture. Every object was specified by its definition (position, shape, color, texture, etc.), its behavior (“is driving”), and by its interrelations with other objects. These interrelations were either hierarchical (“is in”) or not (for groups of objects without hierarchical relation). The description was then stored into an XML document. To obtain the picture description, the blind user worked with a specific browser which went through the objects composing the image and read their information.

The direct use of the physical properties of the sound is another method to represent spatial information. Meijer [77] designed a system (“The Voice”) that uses a time-multiplexed sound to represent a  $64 \times 64$  gray level picture. Every image is processed from left to right and each column is listened to for about 10 milliseconds. Each pixel is associated with a sinusoidal tone, where the frequency corresponds to its vertical position (high frequencies are at the top of the column and low frequencies at the bottom) and the amplitude corresponds to its brightness. Each column of the picture is defined by superimposing the vertical tones. This head-centric coding does not keep a constant pitch for a given object when one nods the head because of elevation change. In addition, interpreting the resulting signal is not obvious and requires extensive training. Capelle et al. [78] proposed the implementation of a crude model of the primary visual system. The implemented device provides two resolution levels corresponding to an artificial central retina and an artificial peripheral retina, as in the real visual system. The auditory representation of an image is similar to that used in “TheVoice” with distinct sinusoidal waves for each pixel in a column and each column being presented sequentially to the listener.

Hollander [79] represented shapes using a “virtual speaker array.” This environment was defined with a virtual auditory spatialization system based on specific head-related transfer function (HRTF) [26]. The auditory environment directly mapped the visual counterpart; a pattern was rendered by a moving sound source that traced in the virtual auditory space the segments belonging to the pattern. Gonzalez-Mora et al. [80] have been working on a prototype for the blind in the Virtual Acoustic Space Project. They have developed a device which captures the form and the volume of the space in front of the blind person’s head and sends this information, in the form of a sound map through

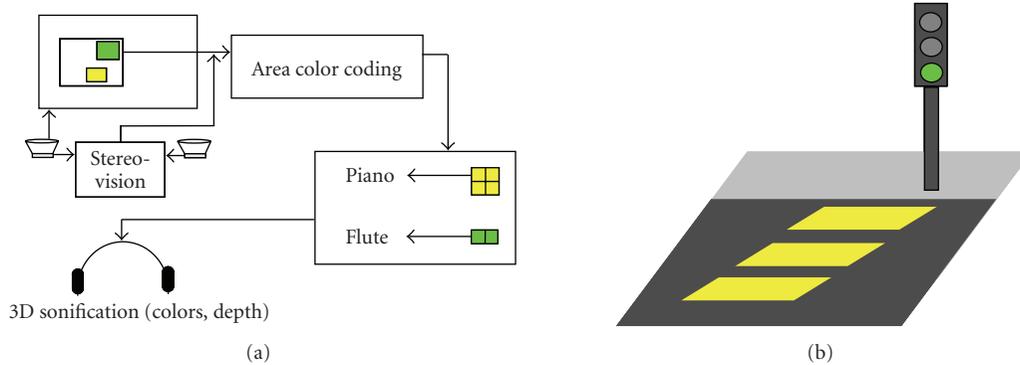


FIGURE 2: Schematic representation of the SeeColor targeted mobility aid. A user points stereo cameras towards the portion of a visual scene that will be sonified. Typical colors, here green for the traffic light and yellow for the crosswalk, are transformed into particular musical instrument sounds: flute for the green pixels, and piano for the yellow ones. These sounds are rendered in a virtual 3D sound space which corresponds to the observed portion of the visual scene. In this sound space, the music from each instrument appears to originate from the corresponding colored pixels location: upper-right for the flute, bottom-center for the piano.

headphones in real time. Their original contribution was to apply the spatialization of sound in the three-dimensional space with the use of HRTFs.

Rather than trying to somehow directly map scene information into audio output, it is also possible to perform some form of image or scene analysis in order to obtain a compact description that can then be spoken to the user. This is typically the case with devices for reading books, such as with the Icare system [81]. Programs that look for textual captions in images also enter in this category; they can be very useful for instance for accessing web pages in which text is often inlaid in images. Similarly, diagram translators allow to describe the content of schematics. Some applications that are more sophisticated in terms of image or video processing often address mobility and life in real, unfamiliar environments. When mobility is concerned, there is need for systems embedded in portable computers such as PDA's. One example that targets unfamiliar environment concerns the design of a face recognition system, where images acquired by a miniature camera located on spectacles are analyzed and then transmitted by a synthetic voice [82].

Concerning developments revolving around navigation, Eddowes and Krahe [83] present an approach for detecting pedestrian traffic lights using color video segmentation and structural pattern recognition. The NAVI (Navigation Assistance for Visually Impaired) system uses a fuzzy-rule-based object identification methodology and outputs results in a stereo headphone (e.g., [84]). In [85, 86], methodologies for the detection of pedestrian crossings and orientation, and for the estimation of their lengths are discussed. A vision-based monitoring application is presented in [87]; it concerns the detection of significant changes from ceiling-mounted cameras in a home environment, in order to generate spoken warnings when appropriate.

A project currently conducted in one of our laboratories (Geneva) and called SeeColor aims at achieving a noninvasive mobility aid for blind users, that uses the auditory pathway to represent in real-time frontal image scenes [88, 89]. Ideally, the targeted system will allow visually impaired or

blind subjects having already seen to build coherent mental images of their environment. Typical colored objects (signposts, mailboxes, bus stops, cars, buildings, sky, trees, etc.) will be represented by sound sources in a three-dimensional sound space that reflects the spatial position of the objects (see Figure 2). Targeted applications are the search for objects that are of particular use for blind users, the manipulation of objects, and the navigation in an unknown environment. SeeColor presents a novel aspect. Pixel colors are encoded by musical instrument sounds, in order to emphasize colored objects and textures that will contribute to build consistent mental images of the environment. Secondly, object depth is (currently) encoded by signal time length with four possible values corresponding to four depth ranges. In terms of image and video processing, images coming from stereo cameras are processed in order to decrease the number of colors and retain only the most significant ones. Work is underway concerning the extraction of intrinsic color properties, in order to discard as much as possible the effect of the illuminants. Another aspect under investigation concerns the determination of salient regions, both spatial and in depth, to be able to suggest a user where to focus attention [90]. Experiments have been conducted first to demonstrate the ability to learn associations between colors and musical instrument sounds. The ability to locate and associate objects of similar colors has been validated with 15 participants that were asked to make pairs with socks of different colors. The current prototype is now being tested as a mobility aid, where a user has to follow a line painted on the ground in an outdoor setting (see Figure 3); real-time sonification combined with distance information obtained from the stereo cameras allows quite accurate user displacement.

## 5. AUDITORY/HAPTIC ENCODING FOR VISION SUBSTITUTION

In view of the limitations of the auditory or of the haptic channels taken independently, it makes sense to combine them in order to design auditory/haptic vision substitution



FIGURE 3: A blindfolded experiment participant having a head mounted camera and following a red serpentine line with the See-Color interface. A video showing this experiment is available for download from <http://cvml.unige.ch/>.

systems. The first multimodal system for presenting graphical information to blind users was the Nomad [91], where a touch sensitive tablet was connected to a synthetic voice generator. Parkes [92] presents and discusses a suite of programs and integrated hardware called TAGW, where TAG stands for Tactile Audio Graphics.

Systems with similar functionalities allowing the rendering of diagrams were also realized for instance by [73, 93] with emphasis on hierarchical auditory navigation. In these systems, the graphical information to render has to be manually prepared beforehand in order to associate particular vocal information with image regions. Commercial tactile tablets with auditory output exist, such as the T3 tactile tablet from the Royal National College for the Blind, UK [94]. The T3 is routinely used in schools for visually handicapped pupils, for instance to allow access to a world encyclopaedia.

The possibility to render more complex information has also been investigated. Kawai and Tomita [95] describe a system that uses stereo vision to acquire 3D objects and render them using a  $16 \times 16$  raising pins display. Synthetic voice is added to provide more information regarding the objects that are presented. Grabowski and Barner [96] extended the system developed by Fritz by adding sonification to the haptic representation. In this approach, the haptic component was used to represent topological properties (size, position) while sonification mapped purely visual characteristics such as colors or textures. More recently, [97], a framework has been developed for generating haptic representations, called force fields, of scenes captured through a simple camera. The

advantage of this approach lies in the fact that the force fields, after being generated, can be stored and processed independently from their source as an individual means of scene representation. The framework in [97] has been used with videos of 3D map models and can be also used with aerial videos for the potential generation of urban force fields. The resulting force fields are processed using either the Phantom Desktop or the CyberGrasp haptic device.

An auditory-haptic system that uses force-feedback devices complemented by auditory information has been designed by [22, 98, 99]. In a first phase, a sighted person has to prepare an image to be rendered by sketching it and associating auditory information to key elements of the drawing. This phase should ultimately be made automatic through the use of image segmentation methods, but this had not been fully implemented as the project concentrated on the rendering aspects and on evaluation. Associated auditory cues differed depending on whether the part to sonify was a contour or a surface. In case of surfaces, the blind user obtained auditory feedback when crossing the object and/or during the whole time he/she pointed to the object surface. Auditory cues were either tones whose pitch depended on the touched object, or spoken words. In addition, haptic feedback describing the object surface was simulated using either a friction or a textural effect. Contours were rendered using kinesthetic feedback, by a virtual force of fixture based on a virtual spring that attracted the mouse cursor towards the contour (see Figure 4). Experiments were first conducted by a Logitech WingMan Force-Feedback mouse. Its working space was found too limited in space (i.e.,  $2.5 \text{ cm} \times 2 \text{ cm}$ ) confirming the assumption of [100]; a specific force-feedback pointing device was thus built, providing a  $11.5 \text{ cm} \times 8 \text{ cm}$  workspace.

In [97], a very promising approach has been presented for the auditory-haptic representation of conventional 2D maps. A series of signal processing algorithms is applied on the map image for extracting the structure information of the map, that is, streets, buildings, and so on, and the symbolic information, that is, street names, special symbols, crossroads, and so on. The extracted structure information is displayed using a grooved line map that is perceived using the Phantom haptic device. The generated haptic map is then augmented with all the symbolic information that is either displayed using speech synthesis for the case of street names, or using haptic interaction features, like friction and haptic texturing. For example, higher friction values are set for the crossroads, while haptic texturing is used to distinguish between special symbols of the map, like hospitals, and so on. During run time, the user interacts with the grooved line map and whenever a special interest point is reached, the corresponding haptic or auditory information is displayed.

In [101], an agent-based system that supports multimodal interaction for providing educational tools for visually handicapped children is described. Interaction modalities are auditory (vocal and nonvocal) and haptic; the haptic interaction is accomplished using the PHANTOM manipulator. A simulation application allows children to explore natural astronomical phenomena, for instance to navigate through virtual planets. Regarding mobility, Coughlan and

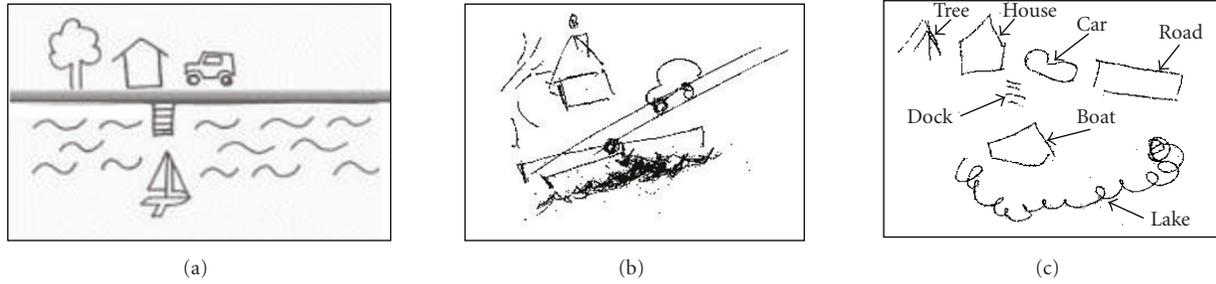


FIGURE 4: Audiotactile rendition of graphs [99]. From left to right: original figure; figure drawn after audiohaptic exploration by a late blind participant; figure drawn by a congenitally blind participant. This illustrates the difference in reconstructed mental images according to the age of appearance of blindness.

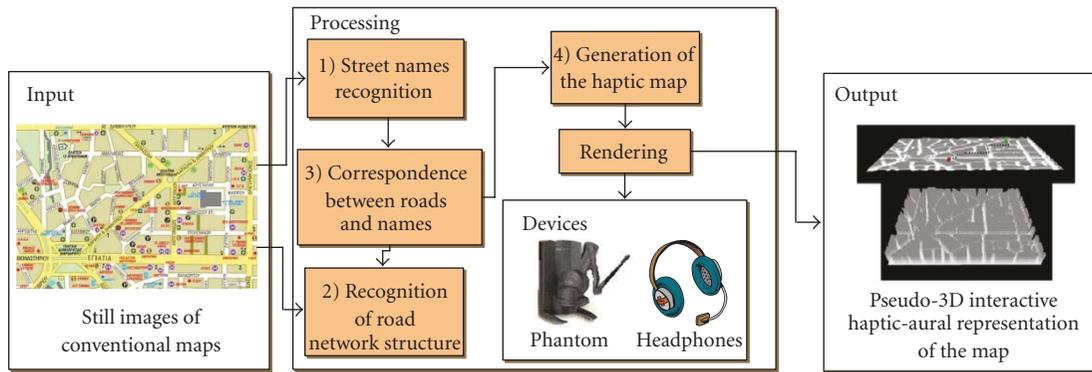


FIGURE 5: Block diagram of the module for the generation of pseudo-3D interactive haptic-aural representations of conventional 2D maps.

Shen [102] and Coughlan et al. [103] address the needs of blind wheelchair users. Their system uses stereo cameras in order to build an environment map. They have also developed specific algorithms to estimate the position and orientation of pedestrian crossings. It is planned to transmit information to the user using synthetic speech, audible tones, as well as tactile feedback.

## 6. CONCLUSIONS

As can be seen from the references, research on vision substitution devices has been active for over a century. Systems aiming at totally replacing the sense of sight for blind persons can be categorized according to the alternate modality that is used to convey the visual information: haptic (tactile and/or kinesthetic), auditory, and auditory/haptic. The use of several modalities is relatively recent, and this trend will necessarily increase since there is a clear benefit in exploiting all possible interaction channels.

The fact that these modalities are of a rather sequential nature implies a fundamental limitation to all visual aids since vision is essentially parallel. A given modality requires some specific preparation of the information. The auditory channel processes an audio signal that is sequential in time, but also allows for some form of parallel processing of the various sound sources composing the stimulus. This “sequential-parallel” capability is for instance used in the SeeColor Project described above: a user sequentially fo-

cuses on various portions of a scene, and each portion is mapped into several simultaneous sound sources. The haptic modality should provide for both global and local analyses; although rather sequential in nature, some form of global parallel exploration is possible when using more than a finger.

For a long time, image/video processing (if any) has remained fairly simple. In many cases, images are prepared manually before being presented to the system. Otherwise, image/video processing can consist of simple thresholding operations, of image simplification techniques based on denoising and contours segmentation. Region segmentation is used for instance to allow region filling with predefined textures. Specific image processing techniques such as contrast enhancement, magnification, and image remapping are used for low-vision aids where the disability to compensate is well characterized spatially or in the frequency domain. There is now a clear trend to use the most recent scene analysis techniques for static images and videos. Object recognition and video data interpretation are performed in order to be able to describe the semantic content of a scene. One reason for this increasing use of fairly involved methods, besides their maturation, is the possibility to embed complex algorithms in portable computers with high processing capabilities.

It is a fact that research in vision replacement does benefit more and more from progress made in computer vision, video and image analysis. Many other issues must however be solved. In terms of human-computer interaction, there

is need to better adapt to user needs in terms of ergonomics and ease of interaction. Attention has to be paid to the appearance of systems to make their use acceptable in public environments (although nowadays wearing “funny looking” devices is not as critical as it was in the 1970s). Regarding evaluation, it is not that easy to find potential users interested in participating in experiments, especially knowing that the devices they are testing most likely will not make it to the market. Not to be neglected is the economic aspect. It is true that the number of totally blind persons is large in absolute numbers, and will increase in relative numbers due to the ageing of the population, but the vast majority of sightless persons cannot easily afford to buy expensive apparatus. Governments therefore should come into play, by providing direct subsidies to those in need as well as funding for research in this area (which is the case now as for instance the 6th and 7th European research programs include such topics).

In conclusion, it is felt that with the current possibilities of miniaturization of wearable devices, the advent of more sophisticated computer vision and video processing techniques, the increase in public funding, more and more visual substitution devices will appear in the future, and very importantly will gain acceptance amongst the potential users.

## ACKNOWLEDGMENTS

This work is supported by the Similar IST Network of Excellence (FP6-507609). T. Pun, P. Roth, and G. Bologna gratefully acknowledge the support of the Swiss Hasler Foundation and of the Swiss “Association pour le bien des aveugles et amblyopes,” as well as the help at various stages of their projects from André Assimacopoulos, Simone Berchtold, Denis Page, and of Professors F. de Coulon (retired) and A. Bullinger (retired) for having helped a long time ago one of the authors (T. Pun) on this fascinating and hopefully useful research topic. Thanks also to many blind persons who have helped us along the years, in particular, Marie-Pierre Assimacopoulos, Alain Barrillier, Julien Conti, and Céline Moret.

## REFERENCES

- [1] World Health Organization, “Magnitude and causes of visual impairment,” Fact Sheet no. 282, November 2004, <http://www.who.int/mediacentre/factsheets/fs282/en/>.
- [2] R. W. Massof and D. L. Rickman, “Obstacles encountered in the development of the low vision enhancement system,” *Optometry and Vision Science*, vol. 69, no. 1, pp. 32–41, 1992.
- [3] E. Peli, L. E. Arend, and G. T. Timberlake, “Computerized image enhancement for visually impaired people: new technology, new possibilities,” *Journal of Visual Impairment & Blindness*, vol. 80, no. 7, pp. 849–854, 1986.
- [4] E. Peli, R. B. Goldstein, G. M. Young, C. L. Trempe, and S. M. Buzney, “Image enhancement for the visually impaired: simulations and experimental results,” *Investigative Ophthalmology & Visual Science*, vol. 32, no. 8, pp. 2337–2350, 1991.
- [5] M. Alonso Jr., A. Barreto, and J. Gualberto Cremades, “Image pre-compensation to facilitate computer access for users with refractive errors,” in *Proceedings of the 6th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '04)*, pp. 126–132, Atlanta, Ga, USA, October 2004.
- [6] M. Alonso Jr., A. Barreto, J. A. Jacko, and M. Adjouadi, “A multi-domain approach for enhancing text with visual aberrations,” in *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '06)*, pp. 34–39, Portland, Ore, USA, October 2006.
- [7] L. Jefferson and R. Harvey, “Accommodating color blind computer users,” in *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '06)*, pp. 40–47, Portland, Ore, USA, October 2006.
- [8] J. A. Brabyn, “New developments in mobility and orientation aids for the blind,” *IEEE Transactions on Biomedical Engineering*, vol. 29, no. 4, pp. 285–289, 1982.
- [9] J. A. Brabyn, “Developments in electronic aids for the blind and visually impaired,” *IEEE Engineering in Medicine and Biology Magazine*, vol. 4, pp. 33–37, 1985.
- [10] J. D. Leventhal, M. M. Uslan, and E. M. Schreier, “A review of technology related publications,” *Journal of Visual Impairment & Blindness*, vol. 84, pp. 127–132, 1990.
- [11] S. M. Kosslyn, *Image and Mind*, Harvard University Press, Cambridge, Mass, USA, 1980.
- [12] M. Carrieras and B. Codina, “Spatial cognition of blind and sighted: visual and amodal hypothesis,” *European Bulletin of Cognitive Psychology*, vol. 12, no. 1, pp. 51–78, 1992.
- [13] J. M. Kennedy, *Drawing and the Blind: Pictures to Touch*, Yale University Press, New Haven, Conn, USA, 1993.
- [14] A. Arditi, J. D. Holtzman, and S. M. Kosslyn, “Mental imagery and sensory experience in congenital blindness,” *Neuropsychologia*, vol. 26, no. 1, pp. 1–12, 1988.
- [15] Y. Hatwell, “Images and non-visual spatial representations in the blind,” in *Non-Visual Human-Computer Interactions*, D. Burger and J.-C. Sperandio, Eds., vol. 228 of *Colloque*, pp. 13–35, INSERM/John Libbey Eurotext, Montrouge, France, 1993.
- [16] R. Passini and G. Proulx, “Way finding without vision: an experiment with congenitally blind people,” *Environment and Behavior*, vol. 20, no. 2, pp. 227–252, 1988.
- [17] S. Ungar, M. Blades, and S. Spencer, “The construction of cognitive maps by children with visual impairments,” in *The Construction of Cognitive Maps*, J. Portugali, Ed., pp. 247–273, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.
- [18] J. Fritz, T. Way, and K. Barner, “Haptic representation of scientific data for visually impaired or blind persons,” in *Proceedings of the 11th Annual Technology and Persons with Disabilities Conference*, Los Angeles, Calif, USA, March 1996.
- [19] E. Hill, J. Rieser, M. Hill, J. Halpin, and R. Halpin, “How persons with visual impairments explore novel spaces: strategies of good and poor performers,” *Journal of Visual Impairment & Blindness*, vol. 87, no. 8, pp. 295–301, 1993.
- [20] H. M. Kamel and J. A. Landay, “A study of blind drawing practice: creating graphical information without the visual channel,” in *Proceedings of the 4th International ACM Conference on Assistive Technologies (ASSETS '00)*, pp. 34–41, Arlington, Va, USA, November 2000.
- [21] H. M. Kamel, P. Roth, and R. R. Sinha, “Graphics and user’s exploration via simple sonics (GUESS): providing interrelational representation of objects in a non-visual environment,” in *Proceedings of the 7th International Conference on Auditory Display (ICAD '01)*, pp. 261–265, Espoo, Finland, July–August 2001.

- [22] P. Roth, "Représentation multimodale d'images digitales dans des systèmes informatiques multimédias pour utilisateurs non-voyants," Ph.D. thesis, Computer Science Department, University of Geneva, Geneva, Switzerland, 2002.
- [23] J. M. Loomis and S. J. Lederman, "Tactual perception," in *Handbook of Perception and Human Performance: Cognitive Processes and Performance*, K. R. Boff, L. Kaufman, and J. P. Thomas, Eds., vol. 2, chapter 31, John Wiley & Sons, New York, NY, USA, 1986.
- [24] S. Millar, *Understanding and Representing Space: Theory and Evidence from Studies with Blind and Sighted Children*, Oxford University Press, Oxford, UK, 1994.
- [25] D. Tzovaras, G. Nikolakis, G. Fergadis, S. Malasiotis, and M. Stavrakis, "Design and implementation of haptic virtual environments for the training of the visually impaired," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 12, no. 2, pp. 266–278, 2004.
- [26] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*, MIT Press, Cambridge, Mass, USA, 1997.
- [27] W. H. Dobelle, D. O. Quest, J. L. Antunes, T. S. Roberts, and J. P. Girvin, "Artificial vision for the blind by electrical stimulation of the visual cortex," *Neurosurgery*, vol. 5, no. 4, pp. 521–527, 1979.
- [28] E. M. Schmidt, M. J. Bak, F. T. Hambrecht, C. V. Kufta, D. K. O'Rourke, and P. Vallabhanath, "Feasibility of a visual prosthesis for the blind based on intracortical microstimulation of the visual cortex," *Brain*, vol. 119, no. 2, pp. 507–522, 1996.
- [29] N. R. Srivastava and P. R. Troyk, "A proposed intracortical visual prosthesis image processing system," in *Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE-EMBS '05)*, pp. 5264–5267, Shanghai, China, September 2005.
- [30] G. Dagnelie and R. W. Massof, "Towards an artificial eye," *IEEE Spectrum*, vol. 33, no. 5, pp. 20–29, 1996.
- [31] M. S. Humayun, R. Freda, I. Fine, et al., "Implanted intraocular retinal prosthesis in six blind subjects," in *Proceedings of the Association for Research in Vision and Ophthalmology (ARVO '05)*, Fort Lauderdale, Fla, USA, May 2005.
- [32] M. Espinosa and E. Ochaita, "Using tactile maps to improve the practical spatial knowledge of adults who are blind," *Journal of Visual Impairment & Blindness*, vol. 92, no. 5, pp. 338–345, 1998.
- [33] J. J. Rieser, "Access to knowledge of spatial structure at novel points of observation," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 15, no. 6, pp. 1157–1165, 1989.
- [34] D. Warren and E. Strelow, *Electronic Spatial Sensing for the Blind*, Martinus Nijhoff, Boston, Mass, USA, 1985.
- [35] R. Easton and B. Bentzen, "The effect of extended acoustic training on spatial updating in adults who are congenitally blind," *Journal of Visual Impairment & Blindness*, vol. 93, no. 7, pp. 405–415, 1999.
- [36] W. Crandall, B. Bentzen, L. Myers, and P. Mitchell, "Transit accessibility improvement through talking signs remote infrared signage, a demonstration and evaluation," Tech. Rep., The Smith-Kettlewell Eye Research Institute, Rehabilitation Engineering Research Center, San Francisco, Calif, USA, 1995.
- [37] R. Golledge, R. Klatzky, and J. Loomis, "Cognitive mapping and way finding by adults without vision," in *The Construction of Cognitive Maps*, J. Portugali, Ed., pp. 215–246, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.
- [38] G. Burdea and P. Coiffet, *Virtual Reality Technology*, John Wiley & Sons, New York, NY, USA, 2003.
- [39] P. J. Standen, D. J. Brown, and J. J. Cromby, "The effective use of virtual environments in the education and rehabilitation of students with intellectual disabilities," *British Journal of Educational Technology*, vol. 32, no. 3, pp. 289–299, 2001.
- [40] M. Schultheis and A. Rizzo, "The application of virtual reality technology for rehabilitation," *Rehabilitation Psychology*, vol. 46, no. 3, pp. 296–311, 2001.
- [41] C. Giess, H. Evers, and H. Meinzer, "Haptic volume rendering in different scenarios of surgical planning," in *Proceedings of the 3rd Phantom Users Group Workshop (PUG '98)*, pp. 19–22, MIT, Cambridge, Mass, USA, October 1998.
- [42] P. Gorman, J. Lieser, W. Murray, R. Haluck, and T. Krummel, "Assessment and validation of force feedback virtual reality based surgical simulator," in *Proceedings of the 3rd Phantom Users Group Workshop (PUG '98)*, MIT, Cambridge, Mass, USA, October 1998.
- [43] G. Jansson, J. Fanger, H. König, and K. Billberger, "Visually impaired persons' use of the phantom for information about texture and 3D form of virtual objects," in *Proceedings of the 3rd Phantom Users Group Workshop*, MIT, Cambridge, Mass, USA, October 1998.
- [44] C. Colwell, H. Petrie, D. Kornbrot, A. Hardwick, and S. Furner, "Haptic virtual reality for blind computer users," in *Proceedings of the 3rd International ACM Conference on Assistive Technologies (ASSETS '98)*, pp. 92–99, Marina del Rey, Calif, USA, April 1998.
- [45] C. Sjöström and K. Rasmus-Gröhn, "The sense of touch provides new computer interaction techniques for disabled people," *Technology and Disability*, vol. 10, no. 1, pp. 45–52, 1999.
- [46] A. Karshmer and C. Bledsoe, "Access to mathematics by blind students: introduction to the special thematic session," in *Proceedings of the 8th International Conference on Computers Helping People with Special Needs (ICCHP '02)*, Linz, Austria, July 2002.
- [47] W. Yu, R. Ramloll, and S. A. Brewster, "Haptic graphs for blind computer users," in *Haptic Human-Computer Interaction*, S. Brewster and R. Murray-Smith, Eds., Springer, Berlin, Germany, 2001.
- [48] P. Parente and G. Bishop, "BATS: the blind audio tactile mapping system," in *Proceedings of the 41st ACM Southeast Regional Conference (ACMSE '03)*, Savannah, Ga, USA, March 2003.
- [49] C. Magnusson, K. Rasmus-Gröhn, C. Sjöström, and H. Danielsson, "Navigation and recognition in complex haptic virtual environments—reports from an extensive study with blind users," in *Proceedings of the Eurohaptics*, Edinburgh, UK, July 2002.
- [50] O. Lahav and D. Mioduser, "Exploration of unknown spaces by people who are blind, using a multisensory virtual environment (MVE)," *Journal of Special Education Technology*, vol. 19, no. 3, pp. 15–24, 2004.
- [51] J. Sánchez and M. Lumbreras, "Virtual environment interaction through 3D audio by blind children," *Cyberpsychology and Behavior*, vol. 2, no. 2, pp. 101–111, 1999.
- [52] S. Semwal and D. Evans-Kamp, "Virtual environments for visually impaired," in *Proceedings of the 2nd International Conference on Virtual Worlds (VW '00)*, vol. 183, pp. 270–285, Paris, France, July 2000.
- [53] C. Grin, "Anoculoscope, appareil à faire voir les aveugles par le sens du toucher," "Description avec dessins photographiques, Paris, chez M. Grin, 6 rue Hippolyte-Lebas", of from Bernard et Cie, 1881, 48 pages. A somehow easier to

- obtain description of this work is: Gallois, "Anoculoscope: instrument pour faire voir les aveugles par le toucher," *Bulletin Le Valentin Haüy*, October 1883.
- [54] T. Pun, "Tactile artificial sight: segmentation of images for scene simplification," *IEEE Transactions on Biomedical Engineering*, vol. 29, no. 4, pp. 293–299, 1982.
- [55] T. P. Way and K. E. Barner, "Automatic visual to tactile translation. I. Human factors, access methods and image manipulation. II. Evaluation of the TACTile image creation system," *IEEE Transactions on Rehabilitation Engineering*, vol. 5, no. 1, pp. 81–105, 1997.
- [56] S. E. Hernandez and K. E. Barner, "Tactile imaging using watershed-based image segmentation," in *Proceedings of the 4th International ACM Conference on Assistive Technologies*, pp. 26–33, Arlington, Va, USA, November 2000.
- [57] S. A. Wall and S. Brewster, "Sensory substitution using tactile pin arrays: human factors, technology and applications," *Signal Processing*, vol. 86, no. 12, pp. 3674–3695, 2006.
- [58] O. Palacz and E. Kurcz, "The usefulness of modified electrophthalm EL-300 designed by Starkiewicz for the blind," Tech. Rep., Department of Pathopsychology of Vision, Medical Academy, Szczecin, Poland, 1977.
- [59] P. Bach-y-Rita, C. C. Collins, F. A. Saunders, B. White, and L. Scadden, "Vision substitution by tactile image projection," *Nature*, vol. 221, no. 5184, pp. 963–964, 1969.
- [60] P. Bach-y-Rita, "Visual information through the skin: a tactile vision substitution system (TVSS)," *Transactions of the American Academy of Ophthalmology and Otolaryngology*, vol. 78, pp. 729–739, 1974.
- [61] Telesensory, <http://www.telesensory.com/>.
- [62] L. H. Goldish and E. Harry, "The optacon: a valuable device for blind persons," *New Outlook for the Blind*, vol. 68, no. 2, pp. 49–56, 1974.
- [63] D. K. Stein, "The Optacon: Past, Present, and Future," National Federation of the Blind (NFB), USA, 1998, <http://www.nfb.org/Images/nfb/Publications/bm/bm98/bm980506.htm>.
- [64] Immersion, Immersion Corp., 2006, <http://www.immersion.com/>.
- [65] Sensable Technology, <http://www.sensable.com/>.
- [66] C. Ramstein and V. Hayward, "The pantograph: a large workspace haptic device for a multi-modal human computer interaction," in *Proceedings of the Conference on Human Factors in Computing Systems (CHI '94)*, pp. 57–58, Boston, Mass, USA, April 1994.
- [67] Logitech, <http://www.logitech.com/>.
- [68] J. P. Fritz and K. E. Barner, "Design of a haptic visualization system for people with visual impairments," *IEEE Transactions on Rehabilitation Engineering*, vol. 7, no. 3, pp. 372–384, 1999.
- [69] G. Nikolakis, K. Moustakas, D. Tzovaras, and M. G. Strintzis, "Haptic representation of images for the blind and the visually impaired," in *Proceedings of the 11th International Conference on Human-Computer Interaction (HCI '05)*, Las Vegas, Nev, USA, July 2005.
- [70] R. Fish, "An audio display for the blind," *IEEE Transactions on Biomedical Engineering*, vol. 23, no. 2, pp. 144–154, 1976.
- [71] L. Kay, "A sonar aid to enhance spatial perception of the blind: engineering design and evaluation," *Radio and Electronic Engineer*, vol. 44, no. 11, pp. 605–627, 1974.
- [72] L. A. Scadden, "Blindness in the information age: equality or irony?" *Journal of Visual Impairment & Blindness*, vol. 78, no. 9, pp. 394–400, 1984.
- [73] A. R. Kennel, "Audiograf: a diagram-reader for the blind," in *Proceedings of the 2nd ACM Conference on Assistive Technologies (ASSETS '96)*, pp. 51–56, Vancouver, BC, Canada, April 1996.
- [74] D. J. Bennett, "Effects of navigation and position on task when presenting diagrams to blind people using sound," in *Diagrammatic Representation and Inference*, vol. 2317 of *Springer Lecture Notes in Artificial Intelligence*, pp. 161–175, Springer, Berlin, Germany, 2002.
- [75] A. King, P. Blenkhorn, D. Crombie, S. Dijkstra, G. Evans, and J. Wood, "Presenting UML software engineering diagrams to blind people," in *Proceedings of the 9th International Conference on Computers Helping People with Special Needs (IC-CHP '04)*, vol. 3118 of *Lecture Notes in Computer Science*, pp. 522–529, Springer, Paris, France, July 2004.
- [76] Z. Mikovec and P. Slavik, "Perception of pictures without graphical interface," in *Proceedings of the 5th ERCIM Workshop on User Interfaces for All (UI4ALL '99)*, Dagstuhl, Germany, November–December 1999.
- [77] P. B. L. Meijer, "An experimental system for auditory image representations," *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 2, pp. 112–121, 1992.
- [78] C. Capelle, C. Trullemans, P. Arno, and C. Veraart, "A real time experimental prototype for enhancement of vision rehabilitation using auditory substitution," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 10, pp. 1279–1293, 1998.
- [79] A. J. Hollander, "An exploration of virtual auditory shape perception," M.S. thesis, University of Washington, Seattle, Wash, USA, 1994.
- [80] J. L. Gonzalez-Mora, A. Rodriguez-Hernandez, L. F. Rodriguez-Ramos, L. Dfaz-Saco, and N. Sosa, "Development of a new space perception system for blind people, based on the creation of a virtual acoustic space," in *Proceedings of the International Work-Conference on Artificial and Natural Neural Networks (IWANN '99)*, vol. 2, pp. 321–330, Alicante, Spain, June 1999.
- [81] T. Hedgpeth, M. Rush PE, V. Iyer, J. Black, M. Donderler, and S. Panchanathan, "iCare-reader: a truly portable reading device for the blind," in *Proceedings of the 9th Accessing Higher Grounds Conference Accessing Media, Web and Technology*, Boulder, Colo, USA, November 2006.
- [82] S. Krishna, G. Little, J. Black, and S. Panchanathan, "A wearable face recognition system for individuals with visual impairments," in *Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '05)*, pp. 106–113, Baltimore, Md, USA, October 2005.
- [83] D. M. Eddowes and J. L. Krahe, "Pedestrian traffic lights recognition in a scene using a PDA," in *Proceedings of the 4th IASTED International Conference on Visualization, Imaging, and Image Processing (VIIP '04)*, Marbella, Spain, September 2004.
- [84] R. Nagarajan, G. Sainarayanan, S. Yaacob, and R. R. Porle, "Fuzzy-rule-based object identification methodology for NAVI system," *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 14, pp. 2260–2267, 2005.
- [85] M. S. Uddin and T. Shioyama, "Detection of pedestrian crossing and measurement of crossing length—an image-based navigational aid for blind people," in *Proceedings of the 8th IEEE Conference on Intelligent Transportation Systems (ITSC '05)*, pp. 331–336, Vienna, Austria, September 2005.
- [86] T. Shioyama, "Computer vision based travel aid for the blind crossing roads," in *Proceedings of the 8th International Conference on Advanced Concepts for Intelligent Vision Systems*

- (ACIVS '06), vol. 4179 of *Lecture Notes in Computer Science*, pp. 966–977, Antwerp, Belgium, September 2006.
- [87] J. A. Martinez-Alarcon and S. J. McKenna, ““Is it as I left it?” A computer vision aid for the blind,” in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC '04)*, vol. 7, pp. 6439–6444, Hague, The Netherlands, October 2004.
- [88] G. Bologna and M. Vinckenbosch, “Eye tracking in coloured image scenes represented by ambisonic fields of musical instrument sounds,” in *Proceedings of the 1st International Work-Conference on the Interplay between Natural and Artificial Computation (IWINAC '05)*, pp. 327–333, Las Palmas, Spain, June 2005.
- [89] G. Bologna, B. Deville, T. Pun, and M. Vinckenbosch, “Transforming 3D coloured pixels into musical instrument notes for vision substitution applications,” *EURASIP Journal on Image and Video Processing*, vol. 2007, Article ID 76204, 14 pages, 2007.
- [90] B. Deville, G. Bologna, M. Vinckenbosch, and T. Pun, “Depth-based detection of salient moving objects in sonified videos for blind users,” in *Proceedings of the 3rd International Conference on Computer Vision Theory and Applications (VIS-APP '08)*, Funchal, Portugal, January 2008.
- [91] D. Parkes, ““Nomad”: an audio-tactile tool for the acquisition, use and management of spatially distributed information by visually impaired people,” in *Proceedings of the 2nd International Symposium on Maps and Graphics for Visually Impaired People*, pp. 24–29, London, UK, April 1988.
- [92] D. N. Parkes, “Tactile audio tools for graphicacy and mobility: “a circle is either a circle or it is not a circle”,” *British Journal of Visual Impairment*, vol. 16, no. 3, pp. 99–104, 1998.
- [93] S. A. Wall and S. Brewster, “Feeling what you hear: tactile feedback for navigation of audio graphs,” in *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, pp. 1123–1132, Montréal, Québec, Canada, April 2006.
- [94] T3 Tactile tablet, Royal National College for the Blind, UK, <http://www.talktab.org/>.
- [95] Y. Kawai and F. Tomita, “Interactive tactile display system: a support system for the visually disabled to recognize 3D objects,” in *Proceedings of the 2nd ACM Conference on Assistive Technologies (ASSETS '96)*, pp. 45–50, Vancouver, BC, Canada, April 1996.
- [96] N. Grabowski and K. E. Barner, “Data visualization methods for the blind using force feedback and sonification,” in *Telemanipulator and Telepresence Technologies V*, vol. 3524 of *Proceedings of SPIE*, pp. 131–139, Boston, Mass, USA, November 1998.
- [97] K. Moustakas, G. Nikolakis, K. Kostopoulos, D. Tzovaras, and M. G. Strintzis, “The force field haptic rendering method: application in haptic access to visual data for the training of the visually impaired,” *IEEE Multimedia Magazine*, vol. 14, no. 1, pp. 62–72, 2007.
- [98] P. Roth, H. Kamel, L. Petrucci, and T. Pun, “A comparison of three nonvisual methods for presenting scientific graphs,” *Journal of Visual Impairment & Blindness*, vol. 96, no. 6, pp. 420–428, 2002.
- [99] P. Roth and T. Pun, “A multimodal system for the non-visual exploration of digital pictures,” in *Proceedings of the 9th ICIP TC13 International Conference on Human-Computer Interaction (INTERACT '03)*, Zürich, Switzerland, September 2003.
- [100] C. Sjöström, “The IT potential of haptics—touch access for people with disabilities,” Licentiate thesis, Certec, Lund University, Lund, Sweden, 1999.
- [101] R. Saarinen, J. Järvi, R. Raisamo, and J. Salo, “Agent-based architecture for implementing multimodal learning environments for visually impaired children,” in *Proceedings of the 7th International Conference on Multimodal Interfaces (ICMI '05)*, pp. 309–316, Trento, Italy, October 2005.
- [102] J. Coughlan and H. Shen, “A fast algorithm for finding cross-walks using figure-ground segmentation,” in *Proceedings of the 2nd Workshop on Applications of Computer Vision, in Conjunction with the European Conference on Computer Vision (ECCV '06)*, Graz, Austria, May 2006.
- [103] J. Coughlan, R. Manduchi, and H. Shen, “Computer vision-based terrain sensors for blind wheelchair users,” in *Proceedings of the 10th International Conference on Computers Helping People with Special Needs (ICCHP '06)*, Linz, Austria, July 2006.