7 Multimodal Feedback in HCI: Haptics, Non-Speech Audio, and Their Applications

Euan Freeman, Graham Wilson, Dong-Bach Vo, Alex Ng, Ioannis Politis, Stephen Brewster

Computer interfaces traditionally depend on visual feedback to provide information to users, with large, high-resolution screens the norm. Other sensory modalities, such as haptics and audio, have great potential to enrich the interaction between user and device to enable new types of interaction for new user groups in new contexts. This chapter provides an overview of research in the use of these non-visual modalities for interaction, showing how new output modalities can be used in the user interface to different devices. The modalities that will be discussed include:

Haptics: *tactons* (*vibrotactile feedback*), thermal (warming and cooling feedback), *force feedback*, and deformable devices;

Non-Speech Audio: *auditory icons*, *Earcons*, *musicons*, sonification, and spatial audio output.

One motivation for using multiple modalities in a user interface is that interaction can be distributed across the different senses or control capabilities of the person using it. If one modality is fully utilized or unavailable (e.g., due to sensory or situational impairment), then another can be exploited to ensure the interaction succeeds. For example, when walking and using a mobile phone, a user needs to focus their visual attention on the environment to avoid bumping into other people. A complex visual interface on the phone may make this difficult. However, haptic or audio feedback would allow them to use their phone and navigate the world at the same time.

This chapter does not present background on multisensory perception and multimodal action, but for insights on that topic see Chapter 2. Chapter 3 also specifically discuss multisensory haptic interaction and the process of designing for it. As a complement, this chapter presents a range of applications where multimodal feedback that involves haptics or non-speech audio can provide usability benefits, motivated by Wickens' *Multiple Resources Theory* [Wickens 2002]. The premise of this theory is that tasks can be performed better and with fewer cognitive resources when they are distributed across modalities. For example, when driving, which is a largely visual task, route guidance is better presented through sound rather than a visual display, as that would compete with the driving for visual cognitive resources. Making calls or texting while driving, both manual tasks, would be more difficult to perform compared to voice dialing, as speech and manual input involve different modalities. For user interface design, it is important to distribute different tasks across modalities to ensure the user is not overloaded so that interaction can succeed.

Definitions

For the purposes of this chapter, a user interface with ***multimodal output*** or feedback is capable of using multiple sensory modalities for presenting information to users (sometimes also known as ***intermodal feedback***). ***Multimodal input*** would allow the use of several different forms of input to a system, for example speech and gesture. In the context of this chapter, we focus on non-speech audio and haptic feedback. This is in contrast to ***multimedia output***, where an application including video, animation and images might be considered multimedia but all use the visual modality.

More specifically, ***crossmodal feedback*** provides exactly the same, or redundant information, across different modalities (see Section 7.1.3). For example, the same information (e.g., amplitude) might be presented using a non-speech sound or vibration. This can be beneficial since in one context audio feedback might be inappropriate, so the user interface could present the same information through haptics instead. This is similar to the idea of sensory substitution in user interfaces for people with disabilities where, for example, visual text might be presented as Braille.

Glossary

Amodal attributes are properties that occur in multiple modalities. For example, audio and tactile icons share many of the same design parameters at the signal level, including frequency, intensity, and rhythm.

Auditory icons are caricatures of natural sounds occurring in the real world, used to represent information from a computer interface [Gaver 1986]. One example is the sound of paper being scrunched up when a document is added to the "recycle bin" on a desktop computer.

Crossmodal feedback is when the same information is presented across different sensory modalities, or redundantly. For example, information (e.g., amplitude) can be presented using either audio or haptic modalities.

Cutaneous sensations come from the skin can include vibration, touch, pressure, temperature, and texture [Lederman and Klatzky 1987].

Earcons are structured abstract audio messages, made from rhythmic sequences called motives [Blattner et al. 1989]. Motives are parameterized by audio properties like rhythm, pitch, timbre, register, and sound dynamics.

Force feedback usually involves displays that move, and can push or pull on part of the body. They generally need to be grounded against something—a table, a car chassis, or another part of the user's body—to provide this resistance or motivate force.

Haptic is a term referring to both ***cutaneous*** sensations gained from the skin, also referred to as ***tactile feedback***, and the ***kinesthetic*** sense, which involves internal signals sent from the muscles and tendons about the position and movement of a limb [van Erp et al. 2010, MacLean 2008a].

Intramodal feedback is feedback that presents information on different aspects of the same sensory modality to achieve a goal. For example, vibrotactile and thermal cues could be combined as intramodal haptic feedback, or force feedback and vibration output could be combined to render texture information.

Kinesthetic signals are sent from muscles and tendons. They include force production, body position, limb direction, and joint angle [van Erp et al. 2010].

Musicons (musical icons) are short audio messages constructed from music snippets [McGee-Lennon et al. 2011].

Tactile Feedback comprises devices that render a percept of the ***cutaneous*** sense: for example, using vibration, temperature, texture, or other material properties to encode information. This term is often used interchangeably with more specific types of tactile feedback, e.g., ***vibrotactile feedback*** and ***thermal feedback***.

Tactons (tactile icons) are structured abstract tactile messages that use properties of vibration to encode information [Brewster and Brown 2004].

Thermal Feedback specifically refers to the use of thermal properties (e.g., temperature change) to encode information.

Thermal Icons are structured thermal changes (i.e., change in temperature) that can convey multidimensional information [Wilson et al. 2012].

Vibrotactile Feedback specifically refers to the use of vibration to encode information.

Intramodal feedback provides information on different aspects of the same sensory modality to achieve a particular goal. For example, force feedback and vibration output could be combined to render texture (see Section 7.1.3.3). Combining cues in this way can support richer feedback in a single modality. This chapter's Glossary provides further definitions of terms, and Focus Questions also are available to aid reader comprehension.

Chapter Outline

Section 7.1 presents an overview of HCI research in the haptic and auditory modalities and provides a high-level summary of how each can be effectively and appropriately utilized in a user interface. Section 7.2 gives specific examples of how the benefits of different modalities have been applied to address real-world problems. It is structured under three themes, which address significant challenges in modern HCI: interaction with mobile devices, making interfaces accessible to users with sensory impairments, and interaction in cars.

Tablets, smartphones, and wearable devices are taking over from the desktop PC as the primary computing device for many people. Interaction design for such devices is difficult as they have small or no visual display and correspondingly limited input space. Using non-visual modalities for feedback, and providing input methods that do not require direct touch, can free visual attention to the mobile environment and provide access to more information than can be viewed on a small screen. The use of multiple modalities, particularly non-visual ones, has also been central in making computer interfaces more accessible to people with sensory impairments, as information can be provided through alternative sensory modalities to enable equal access. As computer interfaces have traditionally been visual, most research on improving accessibility has focused on blind and visually impaired users, converting graphical data, text, and user interface components into audio or haptic forms. Non-visual output can also support visually impaired users in everyday tasks. For example, audio and haptics can be used to help visually impaired users find their way safely and successfully.

While Sections 7.1 and 7.2 show how human perception has been utilized to expand the available information channels, the extent to which HCI has leveraged human sensory capacity is still very limited and new technologies allow us to communicate with people in new ways. Therefore, this chapter concludes with perspectives on the future of multimodal HCI.

7.1 Overview of Non-Visual Feedback Modalities

This section presents a summary of research into non-visual feedback, focusing on haptic and audio output. We take a broad rather than deep look at each modality, to give an idea of the many ways each sense can be used to improve interaction. We then look at how these non-visual modalities can be used together to create multimodal, crossmodal, and intramodal feedback.

7.1.1 Haptics

Haptic is a term referring to both ***cutaneous*** sensations gained from the skin, also referred to as ***tactile feedback***, and the ***kinesthetic*** sense, internal signals sent from the muscles and tendons about the position/movement of a limb [van Erp et al. 2010]. Cutaneous sensations include vibration, touch, pressure, temperature, and texture [Lederman and Klatzky 1987]. Kinesthetic signals include force production, body position, limb direction, and joint angle [van Erp et al. 2010], which supports haptic outputs like force feedback (resistive force) and object deformation/hardness [Lederman and Klatzky 1987], as well as haptic inputs like pressure and gesture input. Both the cutaneous and kinesthetic senses have seen extensive research across many different use cases and implementations, in an effort to leverage human perceptual and manipulative abilities. Here we discuss the research conducted on five major topics within haptic interaction, including two well-established fields (***vibrotactile feedback*** and ***force feedback***) and here emerging fields (***thermal feedback***, ***pressure input***, and ***deformable devices***). Table 7.1 shows how these topics relate to the various aspects of the haptic modality identified here.

7.1.1.1 Tactons: Vibrotactile Feedback

Vibration is the most commonly used type of haptic output due to its ubiquitous use in mobile phones, videogame controllers, smart-watches, and activity trackers. In many cases, vibration is simply used to attract attention (e.g., to notify users of an unread text message) or to give feedback about interaction (e.g., confirming the user pressed a button on the touchscreen). However, vibration has several

Table 7.1 **Summary of haptic interaction techniques surveyed**

Interaction Technique	Input/Output	Haptic Sense	Haptic Channels
Tactons (7.1.1.1)	Output	Cutaneous	Vibration
Thermal feedback (7.1.1.2)	Output	Cutaneous	Temperature
Force feedback (7.1.1.3)	Output	Kinesthetic	Resistance
Pressure input (7.1.1.4)	Input	Both	Pressure
Deformable UIs (7.1.1.4)	Input, Output	Kinesthetic	Resistance, deformation, hardness

dynamic properties, which means it can be used for rich and complex information encoding. Structured abstract messages that use properties of vibration to encode information non-visually are called ***tactons*** (tactile icons) [Brewster and Brown 2004]. Brewster and Brown [2004] outlined a design space for Tactons and described seven properties that could be used for encoding information: frequency, amplitude, waveform, duration, rhythmic patterns, body location, and spatiotemporal patterns. They also described techniques for combining Tactons to create more complex compound messages.

Early research on Tactons looked at which properties of vibration were effective for encoding information. Brown et al. [2005] introduced an additional property of vibration ("roughness") and performed the first evaluation of Tacton identification. Their study evaluated two-dimensional Tactons using roughness and rhythm, finding an overall identification rate of 71%. Rhythm was especially effective (93% identification on its own), with roughness (80%) less so. In a following study [Brown et al. 2006], they designed three-dimensional Tactons using spatial location as the third parameter. Identification rate was 48% when three levels of each parameter were used, although this increased to 81% when roughness was reduced to two levels. This finding shows the potential of high-dimensional information encoding using Tactons. Hoggan and Brewster [2007b] investigated methods of creating vibrotactile "roughness," as Brown's earlier findings were disappointing considering the rich potential of using texture to encode information. They compared amplitude modulation (as in Brown et al. [2005]) with the use of frequency and waveform, finding that frequency (81%) and waveform (94%) significantly outperformed amplitude modulation (61%) in terms of identification.

The examples of Tactons discussed so far have been statically presented against the skin, resulting in abstract structured haptic feedback that has little resemblance to familiar tactile cues, like being tickled or prodded. Li et al. [2008] demonstrated that by moving an actuator relative to the skin as it vibrated, users felt like something was tapping or rubbing against them. A similar approach was used by Ion et al. [2015] with "skin drag displays." These move an actuator against the skin, creating different tactile sensations to static Tactons.

The design of Tactons will often depend on the actuators used to present them. The studies discussed before [Brown et al. 2005, 2006, Hoggan and Brewster 2007b] used voice-coil actuators (e.g., Figure 7.1, top), which were driven by audio signals. These actuators support each of the properties discussed before, although frequency is often limited as each voice coil actuator responds best to a limited frequency bandwidth. Many of today's devices use rotating motors or linear actu-

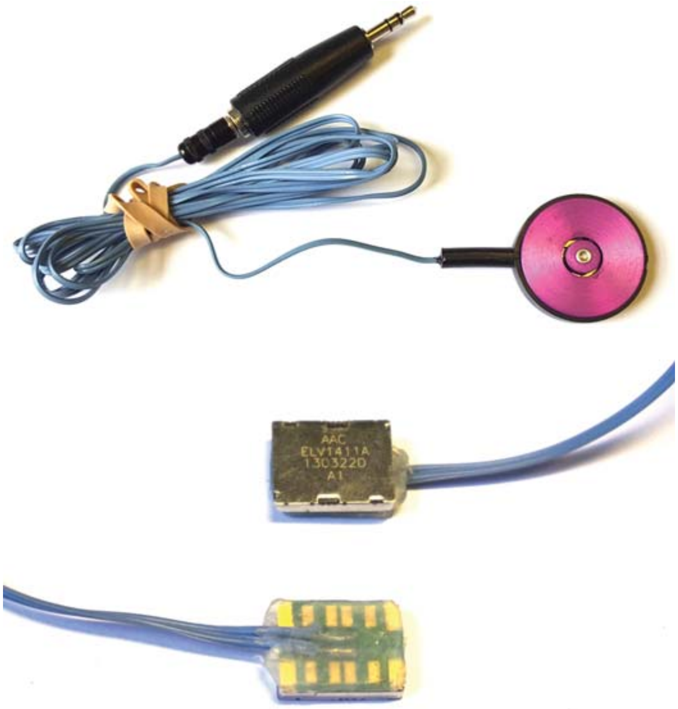

Figure 7.1 **Top: The *EAI C2 Tactor* voice-coil actuator, commonly used in HCI studies for vibrotactile output. Bottom: Two *AAC Technologies ELV-1411A* linear resonant actuators.**

ators (Figure 7.1, bottom), which are simpler. For a review of common vibrotactile displays, see Choi and Kuchenbecker [2013].

Small vibrotactile actuators can be arranged into multi-actuator configurations. For example, six actuators may be placed in a single row or in a 2x3 grid. Such multi-actuator displays can increase the expressiveness of Tactons and vibrotactile feedback by allowing more complex feedback patterns to be delivered. Multi-actuator displays can be used to display fixed spatial patterns and dynamic spatiotemporal patterns. Fixed spatial patterns consist of stimuli from one or more actuators in a fixed configuration, where the location of the stimulus represents information. Dynamic spatiotemporal patterns vary the location of stimuli over time. For example, vibration may "sweep" from left to right or from right to left. Common sites for multi-actuator displays have been the wrist and abdomen. For studies investigating vibrotactile perception and localization at each location, see work by Cholewiak and Collins [2003 2004] as a starting point.

7.1.1.2 Thermal Feedback

Compared to vibrotactile feedback, the thermal sense has been utilized far less in HCI. Thermal perception is an integral part of the cutaneous sense and inherently conveys information about objects (e.g., warmth indicates life) and the environment (e.g., cold indicates danger). It also has inherent links to social (e.g., physical closeness) and emotional (e.g., "warm and loving") phenomena, providing unique opportunities for feedback design. Research in HCI initially looked at what thermal changes are reliably perceivable in different interaction scenarios (i.e., indoors, outdoors, walking and wearing different clothes), to identify what changes and sensations can be used to create feedback (e.g., Wilson et al. [2011]). The research discussed in this section has most often used Peltier devices [Sines and Das 1999] to provide thermal stimulation directly to the skin, as they are available in different sizes, for different devices/use cases, and the exposed surface can be both warmed and cooled. Figure 7.2 shows two Peltier devices used by Wilson et al. [2011 2012 2013 2015].

Thermal feedback has been utilized in a similar way to Tactons [Brown et al. 2006], using structured thermal changes called ***thermal icons***, to convey multidimensional information [Wilson et al. 2012]. Two-dimensional (direction of change and subjective intensity) thermal icons could be identified with 83% accuracy when

Figure 7.2 **Two Peltier modules on black heat sinks, used for thermal stimulation.** The white modules are placed in contact with the skin (e.g., against the palm of the hand).

Figure 7.3 **Screenshot from a video describing how thermal feedback might be used in several application scenarios.** Link to video of slides (no audio): http://doi.org/10.1145/2702123.2702219. (From Wilson et al. [2015])

sitting indoors (97% for the direction, 85% for intensity) [Wilson et al. 2012], but accuracy dropped to 65% when sitting/walking outdoors (96% for direction, 73% for intensity) [Wilson et al. 2013]. While outdoor environments influenced identification, walking had no significant impact. Figure 7.3 gives examples of how thermal feedback might be used to enhance interaction.

7.1.1.3 Force Feedback

When users interact with a physical input device, ***force feedback*** can be given through resistance applied against their movements (resistive) or with them (attractive). Resistive force may be applied to prevent the pointer from moving in certain directions, for example. Attractive force may be applied to guide movements, such as nudging users' input towards targets, for example. One of the earliest devices to provide force feedback was a haptic mouse [Akamatsu and Sato 1994], which could create friction when moved. More advanced devices would follow, capable of applying mechanical resistance against users' movements, in 3D as well as 2D space. For example, one of the most commonly studied force feedback devices is the *SensAble PHANTOM* [Massie and Salisbury 1994], a six degree-of-freedom pointing device that can apply resistance as users move an attached stylus (or thimble placed over their finger), as in Figure 7.4. Recent research has investigated force feedback for non-contact interactions. For example, *FingerFlux* [Weiss et al 2011] could apply attractive or resistive forces against a magnet attached to the fingertip

Figure 7.4 **A SensAble PHANTOM Omni (now Geomagic Touch) with pen attachment. (From McGookin and Brewster [2006])**

as it moved over a flat surface. Others have used ultrasound acoustic radiation pressure to apply a weak resistive force against the skin [Iwamoto et al. 2008, Carter et al. 2013].

Force feedback devices can be used to create more haptic sensations than just being "nudged." For example, they can make virtual objects feel deformable, and can create damping effects on virtual pointers. More sophisticated ***haptic rendering*** [Srinivasan and Basdogan 1997] can use force feedback to create richer haptic effects, like texture. One use of such haptic rendering has been the enhancement of graphical user interfaces with haptic effects. Oakley et al. [2000] described four haptic effects for augmenting buttons in a pointing interface: (1) ***textured surfaces*** allowed users to feel when the pointer is positioned over a button; (2) ***friction*** dampened pointer movements over buttons; (3) ***recesses*** "trapped" the pointer on the button, requiring sufficient velocity to escape; and (4) ***gravity wells*** snapped the pointer towards the middle of a button, helping users stay on the target. For pointing tasks, they found that recesses and gravity wells could reduce the number of errors made. Texture and friction performed poorly because they affected users' pointer movements. Force feedback has also been used to improve the accessibility of graphical data, particularly for visually impaired people. Section 7.2.1 discusses this application area in more detail.

7.1.1.4 Pressure-Based and Deformable Interaction

Every touch/manual action inherently involves a degree of applied pressure (e.g., touch, grasp, and squeeze) and the extent of applied pressure has a purpose or meaning. Therefore, pressure input can enhance touch interaction by allowing users to handle devices in meaningful ways through the application of pressure. For example, McLachlan et al. [2014] recently investigated how a hand grasping a tablet could be used as part of the interaction while the other hand touched the screen (Figure 7.5). Deformable devices are a step towards greater realism, allowing users to provide input by manipulating the physical properties of the device they are interacting with and getting feedback about the nature of the deformation. For example, a deformable user interface may let users alter its shape to change its functionality (e.g., bending a phone so it can be worn as a wrist-watch [Follmer et al. 2012]).

One of the earliest deformable UIs was an elastic cube that users could twist, bend and press. These actions controlled the appearance of a 3D shape shown on a computer screen, manipulating the virtual object in the same way that they manipulate the real object [Murakami and Nakajima 1994]. This style of interaction has inherent visual and haptic feedback, as users can see and feel the effects of their manipulations on the deformable controller. Audio feedback is missing, however, even though this often gives valuable cues when manipulating real objects. For example, a plastic object might creak under pressure as it gets close to snapping. *SoundFlex* [Tahiroğlu 2014] investigated the effects of audio feedback about deformation interaction, looking at real-world sounds like cracking and twanging and also musical cues. They found that audio feedback added valuable cues about the range of deformation (e.g., users could hear when an object was close to 'cracking'), and it allowed users to attribute meaning to their deformations.

Research has moved from deformable controllers to complete systems that are deformable. *Gummi* [Schwesig et al. 2004] introduced the concept of a deformable mobile device with a screen, which could be shaped by users during interaction. They explored interaction techniques based on this concept and found that it was most appropriate for simple tasks like content navigation, rather than more complex tasks like text entry. *PaperPhone* [Lahey et al. 2011] is an entirely deformable device that uses a flexible e-ink screen for display and bend gestures to control a traditional phone interface. Girouard et al. [2015] recently investigated how users might interact with such a phone, and they found that bending the top right corner and squeezing the device were the most popular deformations. Devices do not need to be fully flexible to support deformable interaction. For example, *FlexCase*

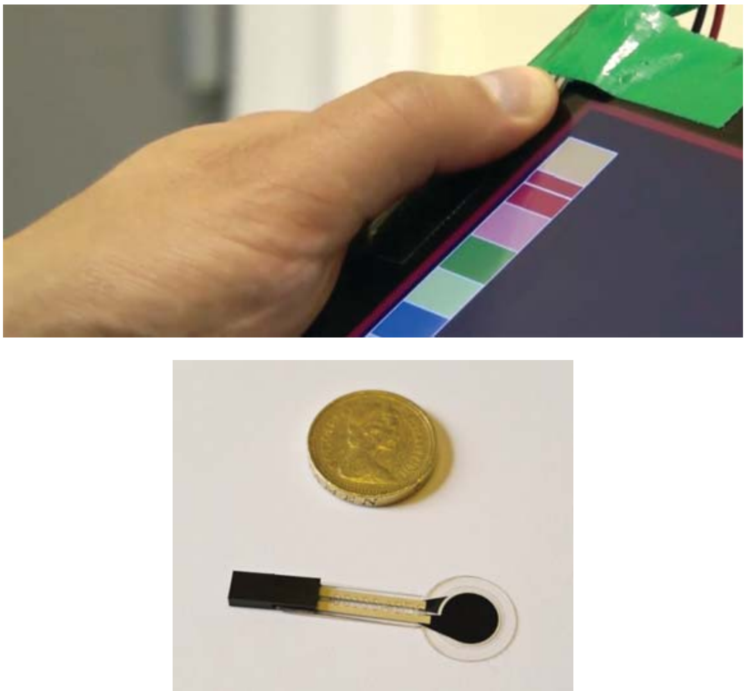

Figure 7.5 **Top: Screenshot from a video that describes a study of pressure as input to a tablet computer.** The still image shows how much pressure is being applied (white line), placed into one of seven levels of pressure. (From [McLachlan et al. 2014]) Bottom: An example of a pressure sensor; this one uses quantum tunneling compound to sense applied pressure. Link to video: http://doi.org/10.1145/2556288.2557260.

[Rendl et al. 2016] was a flexible smartphone cover that users could bend and manipulate to interact with their smartphone.

7.1.2 Non-Speech Audio Feedback: Structured and Representative Sound

Other chapters in this handbook discuss the benefits and challenges of using speech as both an input to a user interface and feedback from it [see Cohen 2017, Katsamanis 2017], Potamianos 2017. While speech can be information-rich and rapid, it also can be cognitively demanding. In some public settings, it likewise can be socially unacceptable as either input or output. This section discusses research on alternative forms of *non-speech* audio feedback, which fall under the general themes of *representative* sound (auditory icons and musicons) or *structured* and abstract sound (sonification, Earcons), compared to the explicit information contained in speech.

7.1.2.1 Auditory Icons

Gaver introduced ***auditory icons*** as a way of representing conceptual objects in a computer system, using sound [Gaver 1986]. He defines them as caricatures of the natural sounds occurring in the real world, using the mapping between a source and the sound it produces to support learning and recall of the meaning of the sound. According to Gaver, the mapping between the sound and the action must be selected carefully. While designers have traditionally used arbitrary mappings between data and their representations, he argues that metaphorical mappings (where representations in the real and virtual worlds are similar) and nomic mappings (physical causation) are generally more meaningful. Thus, using natural sounds rather than manipulating the intrinsic parameters of the sound could greatly improve the learnability of mappings. Like with visual icons, which do not need to be photorealistic representations of real objects, auditory icons do not need to be as accurate as real world sounds. A simple model that captures the essential sound characteristics of an event may be acceptable. Gaver's classic application in this area is the *SonicFinder* [Gaver 1989]. This used auditory icons to represent different aspects of a user interface. Different types of objects in a file browser had different sounds mapped to them: for example, folders had a paper sound, while application icons had metallic sounds.

Two sets of auditory icons have been provided as examples. One set uses the sounds of paper being torn, scrunched up, or thrown away; the other uses the sounds of metallic balls inside a container. There are four "paper" auditory icons.

- http://bit.ly/2nG8rkt is the sound of three pieces of paper being torn.
- http://bit.ly/2noKvpo is the sound of one piece of paper being torn.
- http://bit.ly/2mKrF9g is the sound of paper being scrunched up into a ball.
- http://bit.ly/2mtRU22 is the sound of paper being scrunched up then thrown away.

These auditory icons could be used as audible feedback about operations in a word processing application, for example. The sound of paper being torn might represent content being deleted or "un-done," just like a handwritten note or old draft may be torn up. The sound of paper being scrunched up might represent a document being sent to the wastebasket, just like a discarded piece of paper in the physical world. The latter is a common example in many of today's operating systems: for example, Apple's OSX uses a similar Auditory Icon when a file is moved to the wastebasket.

There are four "metallic ball" auditory icons.

- http://bit.ly/2m1zxGd is the sound of metallic balls being shaken in a container.
- http://bit.ly/2nFXjnC is the sound of three metallic balls falling into a container, one at a time.
- http://bit.ly/2noEWYc is the sound of a handful of metallic balls being dropped into a container.
- http://bit.ly/2muqxFE is the sound of a large handful of metallic balls being dropped into a container.

These auditory icons could be used as audible feedback from a file browser in a desktop operating system, for example. The sound of balls being shaken in a container (Balls_1.mp3) could indicate the number of files in a directory, as the user "picks up" the directory icon and shakes it on screen: the more files there are, the more balls the user would hear rattling in the container. When moving or copying files between directories, the other auditory icons could give feedback about the number of files moved by the operation: the more files the user moved into the directory, the more balls they would hear falling into the container. Contrast this with the simple audio cues used in other desktop operating systems. For example, Apple's OSX plays an abstract synthesized tone when a file has been moved into a new directory. Auditory icons, like the ones discussed here, leverage our familiarity with the physical world to provide additional information: in this case, the number of files being moved.

7.1.2.2 Earcons

Like auditory icons, ***Earcons*** provide audible information about computer objects, operations, or interaction [Blattner et al. 1989, Gaver 1989]. However, unlike auditory icons, which represent a caricature of a realistic sound, Earcons rely on abstract audio representations made from rhythmic sequences called "motives." Each motive is parameterized by rhythm, pitch, timbre, register and sound dynamics [Blattner et al. 1989]. While the rhythm and the pitch set a common basis for a family of motives, combining a specific timbre, register, and dynamic provides a design space to create distinguishable motives from the same family. These motives can also be combined in larger meaningful structures, making Earcons more expressive. However, unlike auditory icons, which are based on recognizable sounds, Earcons require explicit learning.

Brewster et al. [1992] investigated compound and hierarchical Earcons, to see if they were an effective means of communicating complex information. In their experiments, participants had to identify Earcons representing families of icons, menus and combinations of both. Their results show that their more sophisticated Earcon design was significantly more effective than simple beep sounds and was recalled correctly over 80% of the time. They also found that timbre was the most salient feature of the Earcons. Differences in pitch were not recognized as accurately as differences in timbre. Further experiments showed that when presented with a larger structure, participants were able to recall 80% of 27 Earcons arranged in a 4-level hierarchical menu, and up to 97% of 36 Earcons when using compound Earcons [Brewster 1998]. However, a limitation of compound Earcons is that the sound duration increases as the user gets deeper into the menu hierarchy. As an alternative for compound Earcons McGookin and Brewster [2004] investigated the identification of concurrently presented Earcons. They found that increasing the number of concurrently presented Earcons significantly reduced the recognition rate. They suggest introducing a delay of at least 300 ms between successive Earcons, to maximize the chance of successful identification.

Nine example Earcons are provided, which represent menu items in the following menu hierarchy.

Menu 1: Open (http://bit.ly/2lNyiJb), Close (http://bit.ly/2lsf3lz), and Edit (http://bit.ly/2nG8rkt).

Menu 2: Delete (http://bit.ly/2mDXWO7), Create (http://bit.ly/2m4oifr), and Print (http://bit.ly/2lsfyw8).

Menu 3: Copy (http://bit.ly/2lNxx2I), Move (http://bit.ly/2lsaN5C), and Undo (http://bit.ly/2mDUYJ6).

A separate Earcon "family" was created for each menu and all items from the same family have the same timbre (i.e., musical instrument). Menu 1 uses violin sounds, Menu 2 uses electric organ sounds, and Menu 3 uses "fantasy" musical sounds. Individual Earcons vary in terms of their pitch and rhythm. These Earcons can be easily extended. For example, a new item for Menu 1 would use the violin timbre with a distinct rhythmic pattern. A fourth menu could be created by selecting a new musical instrument, e.g., electric guitar or trumpet.

7.1.2.3 Musicons

Earcons typically have no personal meaning to users, which might limit their effectiveness. As an alternative, McGee-Lennon et al. [2011] investigated ***musicons*** (musical icons), short audio clips of music that convey information. By creating musical icons from a user's personal music collection, they hoped to create more recognizable information representations. Their studies suggested that the optimal

Musicon length is 500 ms. McLachlan et al. [2012] found that users prefer musicons that were created from the chorus, or a notable melodic or structural feature, from the songs in their own music library. High identification rates in these studies show the potential of using music to present information to the auditory modality.

The musicons used by McGee-Lennon et al. [2011] have been provided as examples. There are 12 musicons, derived from 4 pieces of music and 3 durations.

- http://bit.ly/2lso6mL, http://bit.ly/2mnk17s, http://bit.ly/2ls5B1L.
- http://bit.ly/2mxs9PA, http://bit.ly/2mntwDk, http://bit.ly/2lhHMZy.
- http://bit.ly/2mE12kT, http://bit.ly/2lPmQuT, http://bit.ly/2lZvc59.
- http://bit.ly/2m4otrf, http://bit.ly/2mxAYZC, http://bit.ly/2l660Ld.

The durations (short, medium and long) correspond to 200 ms, 500 ms, and 1000 ms, respectively. McGee-Lennon et al. [2011] found that the optimal Musicon length was 0.5 s, as this led to the best response time.

7.1.2.4 Sonification of Data

Data are commonly explored using graphical representations. However, visualization techniques are sometimes inadequate for discerning specific features in the data [Kramer 1993]. Sonification, visualization through sound, has the potential to render large sets of high-dimensional data containing variable or temporally complex information. The benefit of sonification is that changes or relationships in data may be easier to hear than to see. In the area of sonification there are several different methods of generating sound from data. Audification, the direct mapping of data samples to audio samples, is the simplest way to make data audible [Kramer 1993], with research showing that audification is perceived to be as efficient as visual graphics for rendering large time-series data sets [Pauletto and Hunt 2005]. One familiar example of audification is the seismogram, for which frequencies are expanded into the audible frequency range [Speeth 1961]. More recently, Alexander et al. [2014] suggested that audification could expose data characteristics that would not be noticeable in visual analysis of very complex data, such as in spectral analysis. A full discussion of audification is outside the scope of this chapter; for more detail, see Alexander et al. [2014] and Dombois and Eckel [2011]. Later in this chapter (Section 7.2.1.1) we discuss the use of sonification to make data accessible to visually impaired users; Figures 7.6 and 7.7 reference videos that demonstrate what this sonification might sound like.

Model-based sonification was introduced by Hermann and Ritter [1999]. They suggest that it could produce more pleasant sounds than audification and could

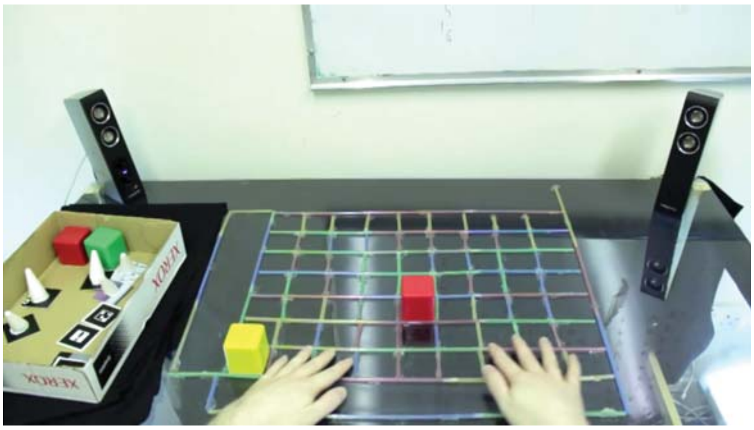

Figure 7.6 **Screenshot from a video demonstrating the *Tangible Graph Builder*.** This combined haptic exploration with sonification to create more accessible ways of interacting with data. Link to video: http://bit.ly/2ntqMlr.

be specifically tailored for task-oriented designs. Model-based sonification may also facilitate learnability, since it uses a limited number of sound parameters. For example, by mapping the physics and dynamic model dimensions of a particle system to sound dimensions, Hermann and Ritter [2004] found it possible to listen to the fluctuation of particle kinetic energy, allowing users to understand the interaction between particles in the system. Another example is the "Shoogle" application, in which users can shake their phone and hear the sound of balls rattling inside, which informs them about the number and size of messages they have received [Williamson et al. 2007]. The impact intensity and pitch of the sounds conveys the mass (size) of each message. For further details, a comprehensive description of model-based sonification techniques is available in Hermann [2011].

A more general and commonly-used approach is parameter-mapping sonification, in which data dimension values are mapped to acoustic parameters of sound. Since sound is multidimensional, this approach is appropriate for rendering multivariate data [Barrass and Kramer 1999]. A common application is auditory graphs, where quantitative changes are mapped to changes in one or more dimensions of sound, such as pitch, panning or timbre [Nees and Walker 2007]. Even though it offers great flexibility, the choice of which acoustic feature to map with the data can affect the effectiveness of the acoustic representation [Walker and Kramer 2005]. Once the acoustic features have been selected, it is essential to consider the interaction of the perceptual dimensions (or orthogonality), as they may distort the perception of the underlying data [Neuhoff et al. 2000]. In addition, other parameters such as polarity (direction of the minimum to the maximum value) or psychophysical scaling (perception of the change between values) must be tested to ensure the success of a parameter-mapping sonification [Walker et al. 2000].

7.1.2.5 Spatial Audio Output

This section has identified several ways of using sound to encode information. These audio feedback techniques can convey further information by using spatial audio output, where the location of the sound relative to the user is meaningful. Spatial audio user interfaces typically use headphones for output, as the stereo earpieces make it easy to position sound relative to the user, especially when they are mobile.

Spatial audio has been used in many types of user interface, with researchers using the location of sound to represent different types of information. Many navigation systems have used spatial audio to let users hear the position of something (e.g., a landmark) relative to their own location. For example, *AudioGPS* [Holland et al. 2002] used spatially encoded Earcons, which increased in frequency as users got closer to a physical landmark. Users could combine the apparent direction of the sound with the temporal frequency to identify where the landmark is and how close they are to it. Blum et al. [2012] used a similar approach, spatially encoding auditory icons and speech output so that visually impaired users could gain a better understanding of their surroundings.

Others have used spatial audio output to give feedback about input, rather than to present information about users' surroundings. For example, Kajastila and Lokki [2013] investigated spatial audio feedback about mid-air menu selection gestures. In their system, the position of sound, relative to the user's head was mapped to the position of menu items, relative to their hand. As users moved their hand to a new menu item, its name was spoken aloud from the direction of the menu item. Their study also investigated visual feedback about the menu selection gestures. They found that an advantage of using spatial audio was that it allowed users to visually focus on their hand and its movements, rather than dividing their attention between the visual feedback and watching what they are doing.

7.1.3 Combined Crossmodal and Intramodal Forms of Feedback

As introduced earlier, multimodal feedback uses multiple sensory modalities to convey information to a user. This section describes research on ***crossmodal feedback*** design, which presents the same information *across* different modalities, and also ***intramodal feedback*** that combines different aspects of within-modality information.

7.1.3.1 Crossmodal Icons

Earlier sections of this chapter introduced Earcons and tactons, structured abstract messages that use sound and vibration to encode information. These non-visual icons share many of the same design parameters, including frequency, intensity and rhythm. Such shared properties can be called ***amodal*** attributes, as they occur in multiple modalities. This shared design space allows the creation of *crossmodal* icons, where the same information can be presented across multiple modalities, either independently or at the same time. An advantage of using crossmodal icons is that it allows an appropriate modality to be chosen based on context. For example, sound may be inappropriate for delivering notifications during meetings, but vibration is subtle and would not disturb others or attract attention. However, not all parameters of Earcons and tactons are suitable for crossmodal display. Table 7.2 shows appropriate properties identified by Hoggan and Brewster [2006, 2007a]. Vibrotactile roughness could be represented in audio using a variety of approaches, e.g., timbre, amplitude modulation, and dissonance. Hoggan and Brewster [2006] found that timbre was the most preferred equivalent of vibrotactile roughness. They also note that intensity can be annoying and has few discriminable levels, so it is not recommended to use as a parameter on its own in crossmodal icons [Hoggan and Brewster 2006, 2007a]. They suggest that frequency is not always appropriate, because of limitations with the vibrotactile actuators they considered. Some contemporary tactile displays do not have such a limited frequency range, however. This highlights the importance of considering technological capabilities when designing user interfaces with multimodal feedback. The range of design properties available depends on the technology available, and there are benefits and trade-offs to consider.

Table 7.2 **Crossmodal mappings between Earcons and tactons**

Earcons	Tactons
Spatial location (3D audio)	Spatial location (actuator position on body)
Rhythm	Rhythm
Timbre	Roughness

Source: [Hoggan and Brewster 2006, 2007a].

In later work [Hoggan and Brewster 2007a], they evaluated the identification rate of three-dimensional crossmodal icons using the parameters shown in Table 7.2. They trained users in one modality, either Earcons or tactons, and tested identification in the other. They also considered the effect of mobility, testing identification while stationary and while mobile. Identification rate ranged from 76.5–85%, suggesting users could successfully identify icons they had learned in another modality, even while walking. Roughness was the worst-performing parameter, consistent with earlier research on tactons [Brown et al. 2005, 2006]. Hoggan et al. [2009] later investigated meaningful mappings between information and crossmodal icon properties, finding that certain audio and tactile parameters were a good fit for certain types of information. They also found that users preferred tactons or crossmodal icons, rather than Earcons, which is feasible whenever device design supports contact with the skin.

7.1.3.2 Combining Visual, Auditory, and Vibrotactile Feedback

Multimodal feedback can be beneficial for warnings while driving, because using multiple sensory channels can quickly and effectively divert attention to important events [Gray et al. 2013, Politis et al. 2015a, 2015b]. It also increases the chance of successfully recognising cues when ambient conditions impair a particular modality. By using more modalities (i.e., trimodal rather than bimodal or unimodal), the perceived urgency of the warnings increases [Politis et al. 2013]. In the presence of critical events, reactions to warnings were quicker with bimodal and trimodal cueing, both in manual driving [Politis et al. 2015b] and autonomous car driving scenarios involving a "hand-over" in which the driver had to resume control of the vehicle [Politis et al. 2015a]. Politis et al. [2015a] also found that bimodal and trimodal warnings that included audio were perceived to be more effective as alerts.

Temporal properties of feedback, like the interval between subsequent pulses or the duration of those pulses, can have an impact on the perceived urgency of messages in the visual, audio, and tactile modalities [Baldwin et al. 2012, van Erp et al. 2015]. When designing alerts of varying urgency, the temporal properties of the feedback should be considered to increase the success of identifying how urgent a message is. Multiple modalities can also be used to make warnings appear more urgent [Politis et al. 2013]. However, in some contexts multimodal presentation can be considered unpleasant or annoying. For example, low priority warnings should be conveyed using fewer modalities, which can be adequately salient with less risk of annoyance [Politis et al. 2013, 2015a]. In this regard, designers must balance salience with user acceptability. As a further design issue, when combining feedback across modalities, the parameters must be perceptually similar enough to be perceived as an integrated unit. For example, auditory pitch and vibrotactile rhythm are perceptually very different, so this combination of modality properties would be likely to confuse users.

We have provided examples of multimodal warnings for drivers, consisting of audio, tactile, and visual signals. The audio warnings can be played through a loudspeaker or headphones. The tactile warnings are intended to drive a C2 tactor attached to the headphone output. The visual warnings are images. These example warnings were used by Politis et al. [2013] in their studies. Three warnings are provided, representing high, medium, and low severity, respectively.

High severity: http://bit.ly/2lPuJQJ (audio), http://bit.ly/2mxpUeY (tactile), http://bit.ly/2lNDPPB (visual).

Medium severity: http://bit.ly/2lNDGM5, http://bit.ly/2lZEM7V, http://bit.ly/2m4pGPb.

Low severity: http://bit.ly/2lPD9rC, http://bit.ly/2lhMO8e, http://bit.ly/2mDTui7.

7.1.3.3 Intramodal Haptic Feedback

As introduced earlier, some modalities have several perceptual aspects that can be presented together as ***intramodal feedback***. For example, a common intramodal haptic feedback combination is force feedback with vibrotactile feedback, which creates textures for virtual objects. Akamatsu and MacKenzie [1996] used this pairing to improve pointing with a mouse. They found that attractive force feedback with an "on-target" tactile stimulus was more effective than a single haptic channel at improving pointer accuracy when selecting small targets.

Others have combined thermal feedback with other haptic stimuli. For example, Gallo et al. [2015] combined thermal feedback with force feedback and judged the comparative stiffness of virtual objects when pressing against the arm. They found that increasing the temperature at the fingertip increased the accuracy of users' stiffness judgments. When designing thermal icons (see Section 7.1.1.2), Wilson et al. [2012] combined thermal feedback, presented to the palm of the hand, with feedback from a vibrotactile actuator, presented to the back of the wrist. The aim was to overcome the limited bandwidth of each individual tactile display. They found that users could identify thermal and vibrotactile messages more accurately via intramodal icons (97%) than purely thermal icons (83%), suggesting thermal and tactile signals can be identified and interpreted simultaneously. These examples demonstrate potential benefits of targeting multiple channels of the same modality.

7.2 Applications of Multimodal Feedback: Accessibility and Mobility

This section gives examples of research that have applied non-visual modalities to enhance interaction. It covers three important and emerging themes in multimodal HCI, where sensory perception is limited: (1) providing accessible interfaces to individuals with visual impairments, (2) supporting interaction with small handheld devices, and (3) presenting information from in-car interfaces while driving. Multimodal feedback is particularly relevant for these topics, because it can be designed to overcome physical or situational impairments.

7.2.1 Multimodal Accessibility

Computer interfaces primarily depend on visual feedback, so the use of multiple non-visual modalities is important for making interfaces accessible to people with visual impairments via "sensory substitution," or presenting information commonly received from one modality via an alternative sense. This section presents research on force-feedback, vibration, and sound to present otherwise visual information, including graphical/tabular data and spatial navigation information.

7.2.1.1 Making Visual Content Accessible through Haptic and Audio Feedback

Graphical data depends on spatial parameters to convey information. The numerical or textual value of graph content can be accessed through sound by visually impaired users (see earlier section on sonification), but the loss of spatial information makes it more difficult to judge relative differences and overall patterns [Lohse 1997]. Researchers have looked at ways of conveying graphical information to visually impaired users, mostly through force feedback devices such as the Logitech WingMan mouse (e.g., Yu et al. [2003]) or the SensAble range of PHANTOM arm devices (e.g., Fritz and Barrier [1999]). In these implementations 2D (in the case of the WingMan) or 3D (for the PHANTOM) graphical charts, such as bar charts or pie charts [Fritz and Barrier 1999, Yu 2002, 2003], can be explored using the mouse- or arm-controlled cursor as an investigative tool (as in Figure 7.7). The devices produce resistive or attractive forces when the cursor contacts the boundaries of chart elements (i.e., bars or planes), to guide movement and convey the spatial properties of the data. However, haptic feedback by itself has only limited benefit. Multimodal systems that also use audio feedback, such as spoken numerical values or sonification [McGookin and Brewster 2006, Yu 2002] support easier navigation around data sets and allow more efficient extraction of information.

As well as providing resistance to user input, the actuated arm on a force feedback device can be moved autonomously to guide the user's hand and let them feel spatial movements or patterns. This method was used by Crossan and

Brewster [2008] to teach visually impaired users 2D trajectories, such as shapes and non-shape patterns. While the force feedback had some success in teaching, the addition of audio feedback representing the arm's position in vertical space (through pitch) and horizontal space (through stereo panning) improved performance, showing the benefits of multimodal feedback in this context.

A similar approach was used to teach visually impaired children about shapes and handwriting [Plimmer et al. 2008, 2011] (Figures 7.8 and 7.9). They attached a pen to a PHANTOM Omni (see Figure 7.4), which mimicked real writing during the task. A teacher drew a shape or letter on a 2D digital screen and then the trajectory was recreated by the PHANTOM device on a horizontal surface. Before training,

Figure 7.7 **Left: Screenshot from a video demonstration of McGookin and Brewster [2006] auditory system for navigating data sets. Link to video: http://bit.ly/2nFZ6sD. Right: Screenshot from a further video demonstration of their work. Link to video: http://bit.ly/2mKtvaa.** These videos show the combined use of force feedback and sonification of tabular data.

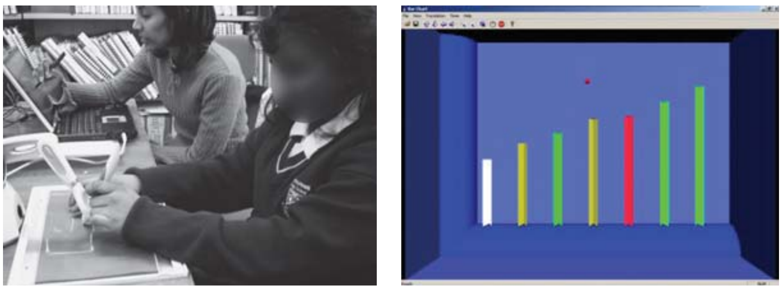

Figure 7.8 **Using force feedback to support visually impaired people with traditionally visual tasks.** Teaching handwriting (left, from Plimmer et al. [2011]) and presenting bar charts (right, from Yu [2002]).

Figure 7.9 **A child using Plimmer et al.'s handwriting system** [2011]. Screen capture from a demonstration video, available at: http://doi.org/10.1145/1993060.1993067.

most children were unable to write basic letters, but all showed improvement after training [Plimmer et al. 2008]. A longitudinal study of the setup added audio feedback to indicate cursor position (stereo panning to indicate horizontal position). It also provided haptic feedback to the writing surface via rubber bands on the upper and lower writing boundaries on lined paper), which the non-PHANTOM hand used to feel and guide the writing [Plimmer et al. 2011]. Letter writing, appropriate spatial positioning, and letter joining all improved over the course of the study, with the children all able to produce a recognizable signature. This section has given examples of how the non-visual modalities we discussed in Section 7.1 can be used to make visual content (e.g., data, shapes, and handwriting) accessible to visually impaired users. Many of these examples also used multimodal output, combining haptic, and audible interactions to enhance interactions with the systems.

7.2.1.2 Haptic and Audio Feedback for Navigation

Multimodal feedback can support navigation for visually impaired people by guiding them to their destination, and informing them about clear and safe paths along the route. Typically, they are guided using spatial haptic or audio directional cues that indicate what direction to move in, a topic that has received substantial research.

A common haptic navigation approach has been to actuate the user's body with belt-like devices. For instance, one haptic belt [van Erp et al. 2005] used vibrotactile actuators to encode the distance and orientation of a reference point. They found that using eight actuators for encoding target location was sufficient for good localization performance, giving a spatial resolution of 45°. By activating two actuators simultaneously, Heuten et al. [2008] were able to improve this resolution to 30°. Flores et al. [2015] compared their wearable haptic guidance system with a speech-based one, finding that the participants navigated faster using the speech system but stayed closer to the intended path when using the haptic system. These examples show how simple spatial vibration output can support visually impaired users by indicating the direction of the next waypoint.

Others have used spatial haptic cues to give information about nearby objects, e.g., obstacles or points of interest. For example, Cardin et al. [2007] informed users about the location of moving obstacles in their path via eight vibrotactile actuators placed along their shoulders. Short (200 ms) bursts with variable intensity were presented once per second by one of eight actuators, positioned left-to-right relative to the obstacle location. Johnson and Higgins [2006] designed a haptic belt system that took visual input from two cameras capturing the user's surroundings and presented tactile stimuli about what the cameras could see. Each section of the visual input was assigned to one of 14 vibrating motors located around the belt. A motor vibrated when an obstacle in the associated section of the image was detected, using intensity of the vibration to encode distance to the obstacle. These examples demonstrate more complex haptic feedback than the navigation research discussed before.

To ensure users do not deviate from the navigation path, which could be dangerous for visually impaired users, research has investigated ways of minimizing path deviation. By interpolating the intensities of two adjacent tactile transducers around a belt, the *Tactile Wayfinder* was capable of providing continuous information about deviations from the path [Heuten et al. 2008]. Marston et al. [2007] investigated whether it was better to tell users if they were on the correct path, or if they were deviating from it. They found that users preferred knowing when they were off course (i.e., feedback given when moving in an incorrect direction). This work shows the importance of investigating the best way of presenting information, as the navigation examples discussed at the start of this section indicated direction of movement, rather than deviation from that direction.

Audio can also be used to convey direction of movement for navigation, or to present information about surroundings. A common approach is to present spatial sound using headphones, an approach we discussed in Section 7.1.2.5. This approach was used by *AudioGPS* [Holland et al. 2002], for example. They used intermittently repeating audio cues like a Geiger counter to encode distance and direction to a point of interest. Audio cues became more frequent as users got closer to the point of interest, and spatialized audio was used so the users could hear where the point was relative to their position and orientation. Blum et al. [2012] used spatialized audio in a similar manner. They combined speech and Auditory Icons to give visually impaired users information about their surroundings. *GpsTunes* [Strachan et al. 2005] continually manipulated a user's music playback to encode direction and distance to a landmark, rather than presenting abstract audio cues (like *AudioGPS* did). Music panned to encode direction, and the volume increased as user's approached the landmark. These works show how the broad range of auditory feedback types, discussed earlier, can be used in similar ways.

A multimodal approach to navigation and information about a user's surroundings was recently demonstrated by Jylhä et al. [2015]. They described a system that supported exploration of urban areas using multimodal audio and haptic feedback from a glove. As users moved around a city, the system informed them of nearby points of interest (e.g., sights, cafes, and shops). It did this using auditory icons and short bursts of vibration. If users showed an interest in a nearby location, the system would then use speech output to tell them more about it. This system demonstrates how multiple output modalities could be used together to provide guidance and spatial information. A multimodal non-visual approach can be beneficial for this purpose. Haptic feedback is more likely to be noticed in noisy urban environments where audio might be obscured by environmental noise. Auditory icons can encode recognizable information without abstract mappings (required by vibration), and speech can present more explicit information. Note that there are trade-offs between speech and non-speech audio, with speech being informative and explicit but more intrusive and time-consuming to listen to. In comparison, non-speech audio (i.e., auditory icons) can be desirable for presenting an overview of landmarks.

7.2.2 Multimodal Interaction with Mobile Devices

The small input and display surface of mobile devices (like phones or watches) means that interaction can be limited and challenging. Interaction can be especially difficult when these devices are used on the move [Barnard et al. 2005, Kjeldskov and Stage 2004] or while carrying other things [Ng et al. 2013, 2014]. However, modern mobile devices have very high-quality audio capabilities along with basic forms of vibration feedback. These multimodal feedback capabilities can be used to overcome some of the problems encountered in everyday interactions on the move. This section gives examples of non-visual feedback for touchscreen interaction and for in-air gestures, an alternative input for small mobile devices. In these cases, the non-visual feedback is presented along with visual output on the screen, creating multimodal feedback.

7.2.2.1 Touchscreen Input and Tactile Feedback

Touchscreens allow designers to develop dynamic user interfaces. Physical buttons can be removed and replaced by their virtual counterparts, although doing so eliminates rich haptic cues. Some designers have restored these rich haptic cues by developing "tactile overlays" that can be added and lie over a touchscreen when needed. These overlays have tactile features, like raised bumps or edges, that mimic the physical features of the on-screen widgets, as in *Touchplates* [Kane et al. 2013]. While such overlays may improve touch input, they are still inflexible like physical buttons and do not support dynamic adaptation.

As an alternative, Poupyrev et al. [2002] explored the use of ambient tactile feedback for tilt-based scrolling tasks on handheld devices. Tactile feedback was used to inform the speed of scrolling in a linear list by presenting a vibrotactile "tap" for every item passed. A user study showed an improvement in overall task completion time when tactile feedback was presented. Hoggan et al. [2008] examined the effectiveness of providing tactile feedback for text entry on a touchscreen mobile phone. Discrete vibrotactile cues were used to indicate whether the finger was on a button, clicking a button, or over the edge of a key on the touchscreen keyboard. Tactile feedback improved the number of phrases entered correctly, compared to typing without tactile feedback, when sitting in a lab setting and in noisy environments such as in a subway. It has also been shown that tactile feedback improves stylus-based text entry on handheld devices in similar mobile settings [Brewster et al. 2007]. This is because the feedback informs users about what is happening in a more noticeable way than through visual feedback alone. For example, the discrete tactile feedback lets users feel that their input was recognized, and lets them feel when they slip off a target. Information about such "slips" would be especially beneficial while users are interacting when mobile. Audio feedback can have similar benefits. For example, Brewster [2002] found that audio feedback improved stylus input accuracy, allowing the creation of even smaller buttons than when visual feedback, alone, is used.

7.2.2.2 Gesture Input with Multimodal Feedback for Mobile Devices

When using small touchscreen devices, users need to be precise to select from targets that are often smaller than their fingers. A way of overcoming this is to move interaction off the screen and into the space *around* the device instead, using gestures in mid-air, rather than touch on the screen, for input. Feedback is important during gesture interaction because it tells users the effects of their actions and it can give them insight into how well their gestures are being sensed. However, mobile devices can only give limited amounts of visual feedback on their small screens.

Figure 7.10 Freeman et al. [2016] **investigated off-screen gesture feedback using three off-screen displays:** LEDs placed around a device (left), tactile displays worn on the hand (right) and the device loudspeaker. (From Freeman et al. [2014])

Freeman et al. [2016] investigated multimodal gesture feedback using three off-screen displays (as in Figure 7.10): LEDs around the device which illuminated surrounding surfaces, sound from the device loudspeaker, and vibration from a device worn on the hand (which they used in earlier work as well [Freeman et al. 2014]). LED output was used to present visual cues, using the layout of the LEDs to give meaningful spatial hints about gesture interaction, like showing users how to move their hand. Audio and tactile output were used to give discrete non-visual feedback about gestures, like a tone or vibration after a successful gesture. These designs leveraged the strengths of each of the modalities: vision has a strong spatial component, making the LED display suitable for presenting spatial cues about gesture movements; and the audio and tactile modalities have strong temporal components, making them suitable for feedback that coincides with users' actions. Audio and tactile feedback presented the same information, with mostly *crossmodal* feedback designs. This meant that audio and tactile feedback could be used together or on their own, when appropriate (e.g., if the user is not wearing a haptics device, audio feedback could be given instead and if users are in a noisy area, tactile feedback could still be perceived). Figure 7.11 gives further examples of this multimodal feedback.

7.2.3 Multimodal Warnings in Cars

Distraction in the car while driving is common, due to secondary in-car activities such as texting, speaking on the phone, or looking at a navigation device [Alm and Nilsson 1994, Salvucci 2001, Summala et al. 1998]. This distraction means that in-car warnings may be missed or may not be noticed in time to have maximum effect. The benefit of using multimodal displays as warnings lies in their ability to attract

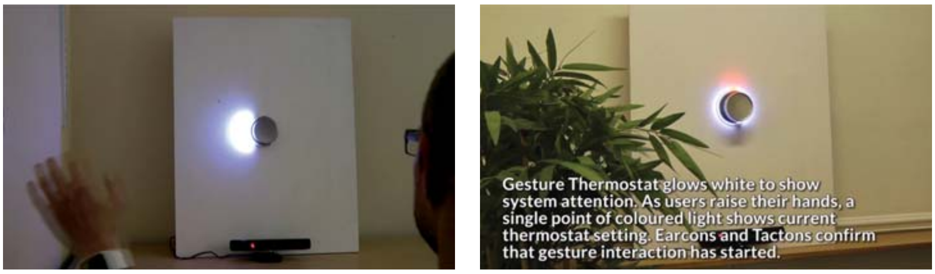

Figure 7.11 **Left: A screenshot from a video demonstration of Freeman et al.'s feedback [2016].** Link to video: http://bit.ly/2munRYB. **This video demonstrates the use of LEDs for feedback, as well as the audio feedback given about gesture input. Right: a further demonstration of similar multimodal feedback being used in a different gesture system Freeman et al. [2015].** Link to video: http://bit.ly/2n13BB1.

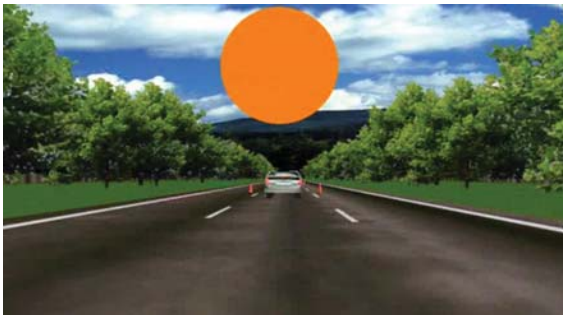

Figure 7.12 **Screenshot from a video demonstration of Politis et al.'s work [2014] on multimodal warnings for cars.** The still image shows their abstract visual warning displayed in a prominent position in the driver's field of view. Link to video: http://doi.org/10.1145/2556288.2556988.

attention when the driver is either distracted or inattentive, and an event on the road requires caution [Ho and Spence 2008]. Modalities that have been used in in-car studies include audio [Ho and Spence 2005], vision [Ablaßmeier et al. 2007], tactile [Ho et al. 2005], and combinations of these [van Erp and van Veen 2001]. Conveying the desired information multimodally has also shown benefit, since the speed and accuracy of reactions improve in this way [Politis et al. 2013, 2014]. Figure 7.12 is a video that demonstrates some multimodal warnings used in this work.

The benefit of multimodal warnings in cars remains relevant even as cars become more automated, taking driving tasks away from the users [Kyriakidis et al.

2014, Meschtscherjakov et al. 2015]. One particularly critical aspect of in-car interaction in autonomous cars involves hand-over of control between the automated car system and driver. Cars are currently not fully autonomous, and are not expected to be so without a transition to partial autonomy first [SAE 2014]. This has motivated research about how to inform the driver of an imminent handover of control and what scenarios on the road would require such a handover [Naujoks et al. 2014, Politis et al. 2015a]. Multimodal warnings are still beneficial in these situations [Politis et al. 2015a]. Indeed, multimodal warnings may be even more useful, because drivers might become more distracted in an autonomous car since they are expected to divert more attention to activities like playing games or email.

7.3 Conclusions and Future Directions

This chapter has discussed a range of existing non-visual feedback techniques for HCI, showing how new feedback methods and technologies can meaningfully change the ways we interact with computers and the ways they can communicate with us. In Section 7.1, we discussed research on haptic feedback and gave examples of how the different perceptual aspects of the haptic modality (the kinesthetic and cutaneous sense) can be targeted with feedback. We also discussed research into non-speech audio feedback, showing many ways of communicating information using representative and structured sounds. Finally, we also introduced the concepts of crossmodal and intramodal feedback and discussed the benefits of using these in user interfaces.

In Section 7.2, we presented three ways non-visual feedback has been used in HCI: to make visual information accessible to visually impaired people, to improve interaction with small handheld devices, and to present information to drivers. These examples demonstrated the benefits of using non-visual feedback to overcome sensory or situational impairments for successful interactions. Many of the feedback techniques discussed in this chapter utilize existing technologies that are well understood, in terms of human perceptual capabilities. As new non-visual feedback technologies emerge, research will be needed to understand their capabilities, human perception of their effects, and their potential applications for HCI. We finish this chapter by discussing two emerging research areas that we think have exciting potential for multimodal HCI: non-contact haptic feedback and shape-shifting interfaces that “act out” against users.

With the primary exceptions of force feedback and deformable devices, computer interfaces have largely remained rigid and passive, detecting an input and producing a corresponding visual, auditory, or haptic response. With improvements in technology, it is now more feasible to have "actuated" devices that can change their physical form, or produce a dynamic physical display that presents information or feedback to the user. An example is 2.5D shape displays: horizontal 2D arrays of small vertically actuating blocks or platforms that individually change height dynamically. These can be used to show information, give feedback, or move other objects [Alexander et al. 2012, Leithinger et al. 2011, Follmer et al. 2013, Robinson et al. 2016]. User-deformable devices like those discussed in Section 7.1.1.4 may also be actuated to change shape automatically in order to provide information or interactive feedback [Ishii et al. 2012], or even to change functions [Yao et al. 2013, Roudaut et al. 2013].

Another emerging area of research is investigating non-contact haptic displays, which can stimulate the haptic modality from a distance. Such haptic displays work by imparting a force upon the user, with sound or air as the delivery mechanism, rather than a device in contact with the skin. The advantage of using such a non-contact haptic display is that users do not have to be instrumented with a device, and do not have to touch something in order to experience the feedback. This could allow haptic feedback in situations where it was previously unavailable. For example, user interfaces to support surgery could use mid-air gesture interactions to avoid the risk of infection and contamination, with non-contact haptics giving feedback about input. Ultrasound haptic displays are an example of an emerging non-contact haptic display. They use an array of ultrasound loudspeakers (as in Figure 7.13) to focus inaudible sound upon a "focal point," which imparts acoustic radiation pressure against the skin that is felt as vibration. This approach was first demonstrated by Iwamoto et al. [2008] and has been refined in recent years. For example, Carter et al. [2013] allowed the creation of multiple mid-air haptic

Figure 7.13 **Ultrasound haptic displays use an array of ultrasound speakers,** which focus sound to create a focused area of acoustic radiation pressure.

focal points, which can be positioned and moved in 3D space above the display. Ultrasound haptic output shares many of the same properties of vibrotactile output (discussed in Section 7.1.1.1), although its perception is not well understood.

With more perceptual research, the HCI community will be able to form a better understanding of how these novel types of technology can be used most effectively to improve interaction. Both are likely to benefit from multimodal feedback—for example, using intramodal haptic feedback to enhance haptic sensations on deformable devices, and combining audio and ultrasound to improve non-contact haptics displays. These and other promising directions remain to be tested in future research.

Focus Questions

7.1. Describe the difference between cutaneous and kinesthetic haptic perception.

7.2. Identify three properties of vibration that are appropriate for tacton design. Then describe a set of nine tactons using these properties. Do you think a user could reliably identify each of your tactons?

7.3. Imagine a smartphone could heat up or cool down to deliver thermal feedback. Describe how you would use "hot" and "cold" when presenting email notifications.

7.4. Describe the difference between auditory icons, Earcons, and musicons.

7.5. Smartphones typically present a short auditory notification when a new email is received. Consider the non-speech audio techniques covered in this chapter, and discuss a new auditory notification design that also encodes the "importance" of a new email and the number of unread emails in the inbox.

7.6. In Section 7.1.3.1 we identified three amodal properties of audio and haptic feedback. Can you think of any other amodal properties of these modalities? Are there any amodal properties that occur between the auditory, haptic, and visual modalities?

7.7. We described research that used vibration to present haptic guidance and navigation instructions. Consider enhancing these methods with thermal feedback: how could you use thermal feedback and vibration to create intramodal haptic guidance instructions? Explain what "hot" and "cold" would represent in your design.

7.8. Touchscreen interaction with mobile devices is compromised when users are walking. If users were to interact using speech commands, they would not have to look at the screen while mobile. Describe how non-visual modalities could be used to give feedback about speech commands while walking—what information might users want to know and how could this be encoded using non-visual feedback?

7.9. We discussed multimodal warnings about in-car events. What modalities would you use to: (a) notify the driver about a new email; (b) inform the driver that their fuel is low; and (c) warn the driver of an imminent collision?

References

M. Ablaßmeier, T. Poitschke, F. Wallhoff, K. Bengler, G. and Rigoll. 2007. Eye gaze studies comparing head-up and head-down displays in vehicles. *IEEE International Conference on Multimedia and Expo*, pp. 2250–2252. DOI: 10.1109/ICME.2007.4285134. 305

M. Akamatsu and I. MacKenzie. 1996. Movement characteristics using a mouse with tactile and force feedback. *International Journal of Human-Computer Interaction,* 483–493. DOI: 10.1006/ijhc.1996.0063. 297

M. Akamatsu and S. Sato, S. 1994. A multi-modal mouse with tactile and force feedback. *International Journal of Human-Computer Studies*, 40(3):443–453. DOI: 10.1016/j.promfg.2015.07.388. 285

J. Alexander, A. Lucero, and S. Subramanian. 2012. Tilt displays: designing display surfaces with multi-axis tilting and actuation. *Proceedings MobileHCI '12*, pp. 161–170. DOI: 0.1145/2371574.2371600. 307

R. L. Alexander, S. O'Modhrain, D. A. Roberts, J. A. Gilbert, and T. H. Zurbuchen. 2014. The bird's ear view of space physics: Audification as a tool for the spectral analysis of time series data. *Journal of Geophysical Research Space Physics* 119(7):5259–5271. DOI: 10.1002/2014JA020025. 292

H. Alm and L. Nilsson. 1994. Changes in driver behaviour as a function of handsfree mobile phones - A simulator study. *Accident Analysis and Prevention* 26(4):441–451. DOI: 10.1016/0001-4575(94)90035-3. 304

C. L. Baldwin, J. L. Eisert, A. Garcia, B. Lewis, S. Pratt, and C. Gonzalez. 2012. Multimodal urgency coding: auditory, visual, and tactile parameters and their impact on perceived urgency. *WORK: A Journal of Prevention, Assessment & Rehabilitation*, 41:3586–3591. DOI: 10.3233/WOR-2012-0669-3586. 296

L. Barnard, J. S. Yi, J. A. Jacko, and A. Sears. 2005. An empirical comparison of use-in-motion evaluation scenarios for mobile computing devices. *International Journal of Human-Computer Studies*, 62(4):487–520. DOI: 10.1016/j.ijhcs.2004.12.002. 302

S. Barrass and G. Kramer. 1999. Using sonification. *Multimedia System*, 7(1):23–31. DOI: 10.1007/s005300050108. 293

M. Blattner, D. Sumikawa, and R. Greenberg. 1989. Earcons and icons: their structure and common design principles. *Human-Computer Interaction*, 4(1):11–44. DOI: 10.1207/s15327051hci0401_1. 279, 290, 613

J. R. Blum, M. Bouchard, and J. R. Cooperstock. 2012. What's around me? Spatialized audio augmented reality for blind users with a smartphone. *Proceedings of the Mobiquitous '12*, pp. 49–62. DOI: 10.1007/s11036-012-0425-8. 294, 302

S. Brewster. 1998. Using nonspeech sounds to provide navigation cues. *ACM Transactions of Computer-Human Interaction*, 5(3):224–259. DOI: 10.1145/292834.292839. 291

S. Brewster. 2002. Overcoming the Lack of Screen Space on Mobile Computers. *Person.al and Ubiquitous Computing* 6(3):188–205. DOI: 10.1007/s007790200019. 303

S. Brewster and L. M. Brown. 2004. Tactons: structured tactile messages for non-visual information display. *Proceedings of the AUIC '04*, pp. 15–23. 279, 282, 631

S. Brewster, F. Chohan, and L. Brown. 2007. Tactile feedback for mobile interactions. *Proceedings of the CHI '07*, pp. 159–162. DOI: 10.1145/1240624.1240649. 303

S. Brewster, D. K. McGookin, and C. A. Miller. 2006. Olfoto: designing a smell-based interaction. *Proceedings of the CHI '06*, pp. 653–662.

S. Brewster, P. C. Wright, and A. D. N. Edwards. 1992. A detailed investigation into the effectiveness of earcons. *Proceedings of the ICAD '92*, pp. 471–498. 290

L. Brown, S. Brewster, and H. C. Purchase. 2005. A first investigation into the effectiveness of tactons. *Proceedings of the WHC '05*, pp. 167–176. 282, 296

L. Brown, S. Brewster, and H. Purchase. 2006. Multidimensional tactons for non-visual information presentation in mobile devices. *Proceedings of the MobileHCI '06*, pp. 231–238. DOI: 10.1145/1152215.1152265. 282, 284, 296

S. Cardin, D. Thalmann, and F. Vexo. 2007. A wearable system for mobility improvement of visually impaired people. *The Visual Computer* 23(2) 109–118. DOI: 10.1007/s00371-006-0032-4. 301

T. Carter, S. A. Seah, B. Long, B. Drinkwater, and S. Subramanian. 2013. UltraHaptics: multi-point mid-air haptic feedback for touch surfaces. *Proceedings of the UIST '13*, pp. 505–514. DOI: 10.1145/2501988.2502018. 286, 307

S. Choi and K. J. Kuchenbecker. 2013. Vibrotactile display: perception, technology, and applications. *Proceedings of the IEEE,* 101(9):2093–2104. DOI: 10.1109/JPROC.2012.2221071. 283

R. W. Cholewiak, J. C. Brill, and A. Schwab. 2004. Vibrotactile localization on the abdomen: effects of place and space. *Perception & Psychophysics,* 66(6):970–987. DOI: 10.3758/BF03194989. 283

R. W. Cholewiak and A. A. Collins. 2003. Vibrotactile localization on the arm: effects of place, space, and age. *Perception & Psychophysics,* 65(6):1058–1077. DOI: 10.3758/BF03194834. 283

P. R. Cohen. 2017. Multimodal Speech and Pen Interfaces. In S. Oviatt, B. Schuller, P. Cohen, D. Sonntag, G. Potamianos, and A. Krüger, editors, *Handbook of Multimodal-Multisensor Interfaces, Volume 1: Foundations, User Modeling, and Common Modality Combinations*. Morgan & Claypool, Williston, VT. 288

A. Crossan and S. Brewster. 2008. Multimodal trajectory playback for teaching shape information and trajectories to visually impaired computer users. *ACM Transactions on Accessible Computing* 1(2):1–34. DOI: 10.1145/1408760.1408766. 299

F. Dombois and G. Eckel. 2011. Audification. In T. Hermann, A. Hunt, and J. G. Neuhoff, editors, *The Sonification Handbook,* pp. 301–324. Logos Publishing House, Berlin. 292

G. Flores, S. Kurniawan, R. Manduchi, E. Martinson, L. M. Morales, and E. A. Sisbot. 2015. Vibrotactile guidance for wayfinding of blind walkers. *IEEE Transactions on Haptics*, 8(3):306–317. DOI: 10.1109/TOH.2015.2409980. 301

S. Follmer, D. Leithinger, A. Olwal, N. Cheng, and H. Ishii. 2012. Jamming user interfaces: programmable particle stiffness and sensing for malleable and shape-changing devices. *Proceedings of the UIST '12*, pp. 519–528. DOI: 10.1145/2380116.2380181. 287

S. Follmer, D. Leithinger, A. Olwal, A. Hogge, and H. Ishii. 2013. inFORM: dynamic physical affordances and constraints through shape and object actuation. *Proceedings of the UIST '13*, pp. 417–426. DOI: 10.1145/2501988.2502032. 307

E. Freeman, S. Brewster, and V. Lantz. 2014. Tactile feedback for above-device gesture interfaces: adding touch to touchless interactions. *Proceedings of the ICMI '14*, pp. 419–426. DOI: 10.1145/2663204.2663280. 304

E. Freeman, S. Brewster, and V. Lantz. 2015. Towards in-air gesture control of household appliances with limited displays. *Proceedings INTERACT '15 Posters in LNCS 9299*, pp. 611–615. DOI: 10.1007/978-3-319-22723-8_73. 305

E. Freeman, S. Brewster, and V. Lantz. 2016. Do that, there: an interaction technique for addressing in-air gesture systems. *Proceedings CHI '16*. DOI: 10.1145/2858036.2858308. 304, 305

J. P. Fritz and K. E. Barrier. 1999. Design of a haptic data visualization system for people with visual impairments. *IEEE Transactions on Rehabilitation Engineering*, 7(3):372–384. 298

S. Gallo, G. Rognini, L. Santos-Carreras, T. Vouga, O. Blanke, and H. Bleuler. 2015. Encoded and crossmodal thermal stimulation through a fingertip-sized haptic display. *Frontiers in Robotics and AI* 2(Oct.):1–12. DOI: 10.3389/frobt.2015.00025. 297

W. Gaver. 1986. Auditory icons: using sound in computer interfaces. *Human-Computer Interaction*, 2(2):167–177. DOI: 10.1207/s15327051hci0202_3. 279, 289, 610

W. Gaver. 1989. The SonicFinder: an interface that uses auditory icons. *Human-Computer Interact.*, 4(1):67–94. DOI: 10.1207/s15327051hci0401_3. 289, 290

A. Girouard, J. Lo, M. Riyadh, F. Daliri, A. K. Eady, and J. Pasquero. 2015. One-handed bend interactions with deformable smartphones. *Proceedings of the CHI '15*, ACM Press, pp. 1509–1518. 287

R. Gray, C. Spence, C. Ho, and H. Tan. 2013. Efficient Multimodal Cuing of Spatial Attention. *Proceedings of the IEEE* 101(9):2113–2122. 296

T. Hermann. 2011. Model-based sonification. In T. Hermann, A. Hunt, and J. G. Neuhoff, editors, *The Sonification Handbook,* pp. 399–427. Logos Publishing House, Berlin. 293

T. Hermann and H. Ritter. 1999. Listen to your data: model-based sonification for data analysis. *International Symposium on Intelligent Multimedia and Distance Education*. pp. 189–194. 292

T. Hermann and H. Ritter. 2004. Sound and meaning in auditory data display. *Proceedings of the IEEE*, 92(4):730–741. DOI: 10.1109/JPROC.2004.825904. 293

W. Heuten, N. Henze, S. Boll, and M. Pielot. 2008. Tactile wayfinder: a non-visual support system for wayfinding. *Proceedings of the NordiCHI '08*, p. 172. DOI: 10.1145/1463160.1463179. 301

C. Ho and C. Spence. 2005. Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. *Journal of Experimental Pyschology: Applied* 11(3):157–174. DOI: 10.1037/1076-898X.11.3.157. 305

C. Ho and C. Spence. 2008. *The multisensory driver: Implications for ergonomic car interface design*. Ashgate Publishing. 305

C. Ho, H. Z. Tan, and C. Spence. 2005. Using spatial vibrotactile cues to direct visual attention in driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(6):397–412. 305

E. Hoggan and S. Brewster. 2006. Crossmodal icons for information display. *CHI '06 Extended Abstracts*, p. 857. DOI: 10.1145/1125451.1125619. 295

E. Hoggan and S. Brewster. 2007a. Designing audio and tactile crossmodal icons for mobile devices. *Proceedings of the ICMI '07*, pp. 162–169. 295, 296

E. Hoggan and S. Brewster. 2007b. New parameters for tacton design. *CHI '07 Extended Abstracts*, pp. 2417–2422. DOI: 10.1145/1240866.1241017. 282

E. Hoggan, S. Brewster, and J. Johnston. 2008. Investigating the effectiveness of tactile feedback for mobile touchscreens. *Proceedings of the CHI '08*, p. 1573. DOI: 10.1145/1357054.1357300. 303

E. Hoggan, R. Raisamo, and S. Brewster. 2009. Mapping information to audio and tactile icons. *Proceedings of the ICMI '09*, pp. 327–334. 296

S. Holland, D. R. Morse, and H. Gedenryd. 2002. AudioGPS: spatial audio navigation with a minimal attention interface. *Person. Ubiquitous Comput.,* 6:253–259. DOI: 10.1007/s007790200025. 294, 301

A. Ion, E. J. Wang, and P. Baudisch. 2015. Skin drag displays: dragging a physical tactor across the user's skin produces a stronger tactile stimulus than vibrotactile. *Proceedings of CHI '15*, ACM Press, pp. 2501–2504. DOI: 10.1145/2702123.2702459. 282

H. Ishii, D. Lakatos, L. Bonanni, and J.-B. Labrune. 2012. Radical atoms: beyond tangible bits, toward transformable materials. *ACM Interactions*, 38–51. DOI: 10.1145/2065327.2065337. 307

T. Iwamoto, M. Tatezono, and H. Shinoda. 2008. Non-contact method for producing tactile sensation using airborne ultrasound. *Proceedings of the EuroHaptics '08*, Springer, pp. 504–513. DOI: 10.1007/978-3-540-69057-3_64. 286, 307

L. A. Johnson and C. M. Higgins. 2006. A navigation aid for the blind using tactile-visual sensory substitution. *2006 International Conference of the IEEE Engineering in Medicine and Biology Society,* pp. 6289–6292. DOI: 10.1109/IEMBS.2006.259473. 301

A. Jylhä, Y-T. Hsieh, V. Orso, S. Andolina, L. Gamberini, and G. Jacucci. 2015. A Wearable Multimodal Interface for Exploring Urban Points of Interest. *Proceedings of the ICMI '15*, pp. 175–182. DOI: 10.1145/2818346.2820763. 302

R. Kajastila and T. Lokki. 2013. Eyes-free interaction with free-hand gestures and auditory menus. *International Journal of Human-Computer interaction*, 71(5):627–640. DOI: 10.1016/j.ijhcs.2012.11.003. 294

S. K. Kane, M. Ringel Morris, and J. O. Wobbrock. 2013. Touchplates: low-cost tactile overlays for visually impaired touch screen users. *Proceedings of the ASSETS '13*, Article 22. DOI: 10.1145/2513383.2513442. 303

A. Katsamanis, 2017. Multimodal Gesture Recognition. In S. Oviatt, B. Schuller, P. Cohen, D. Sonntag, G. Potamianos, and A. Krüger, editors, *Handbook of Multimodal-Multisensor Interfaces, Volume 1: Foundations, User Modeling, and Common Modality Combinations*. Morgan & Claypool, Williston, VT. 288

J. Kjeldskov and J. Stage. 2004. New techniques for usability evaluation of mobile systems. *International Journal Human-Computer Studies,* 60(5–6):599–620. 302

G. Kramer. 1993. *Auditory Display: Sonification, Audification, and Auditory Interfaces*. Perseus Publishing. 292

M. Kyriakidis, R. Happee, and J. C. F. de Winter. 2014. Public opinion on automated driving: Results of an international questionnaire among 5,000 respondents. SSRN 2506579. 305, 306

B. Lahey, A. Girouard, W. Burleson, and R. Vertegaal. 2011. PaperPhone: understanding the use of bend gestures in mobile devices with flexible electronic paper displays. *Proceedings of the CHI '11*, pp. 1303–1312. DOI: 10.1145/1978942.1979136. 287

S. J. Lederman and R. L. Klatzky. 1987. Hand movements: A window into haptic object recognition. *Cognitive Psychology* 19:342–368. 279, 281

D. Leithinger, D. Lakatos, A. DeVincenzi, M. Blackshaw, and H. Ishii. 2011. Direct and Gestural Interaction with Relief: A 2.5D Shape Display. *Proceedings of the UIST '11*, pp. 541–548). DOI: 10.1145/2047196.2047268. 307

K. Li, P. Baudisch, W. G. Griswold, and J. D. Hollan. 2008. Tapping and rubbing: exploring new dimensions of tactile feedback with voice coil motors. *Proceedings the of UIST '08*, ACM Press, p. 181. DOI: 10.1145/1449715.1449744. 282

G. Lohse. 1997. Models of graphical perception. In M. Helander, T. Landauer, and P. Prabhu, editors, *Handbook of Human-Computer Interaction*, pp. 107–135. Elsevier, Amsterdam . 298

K. MacLean. 2008a. Haptic interaction design for everyday interfaces. In M. Carswell, editor, *Reviews of Human Factors and Ergonomics*, pp. 149–194. Human Factors and Ergonomics Society. 279

J. R. Marston, J. M. Loomis, R. L. Klatzky, and R. G. Golledge. 2007. Nonvisual route following with guidance from a simple haptic or auditory display. *Journal of Visual Impairment Blindness,* 101(4):203–211. 301

T. H. Massie and J. K. Salisbury. 1994. The PHANTOM haptic interface: a device for probing virtual objects. *Proceedings of the ASME Dynamic Systems and Control Division*, pp. 295–299. 285

M. McGee-Lennon, M. Wolters, R. McLachlan, S. Brewster, and C. Hall. 2011. Name that tune: musicons as reminders in the home. *Proceedings of the CHI '11*, p. 2803. DOI: 10.1145/1978942.1979357. 279, 291, 292

D. K. McGookin and S. Brewster. 2004. Understanding concurrent Earcons: applying auditory scene analysis principles to concurrent Earcon recognition. *ACM Transactions on Applied Perception,* 1(2):130–155. DOI: 10.1145/1024083.1024087. 291

D. K. McGookin and S. Brewster. 2006. Graph builder: constructing non-visual visualizations. *Proceedings of the HCI'06 Conference on People and Computers XX*, pp. 263–278. DOI: 10.1007/978-1-84628-664-3_20. 286, 298, 299

D.K. McGookin, E. Robertson and S. Brewster. 2010. Clutching at Straws: Using Tangible Interaction to Provide Non-Visual Access to Graphs. In *Proceedings CHI '10*, pp. 1715-1724. ACM Press. DOI: 10.1145/1753326.1753583.

R. McLachlan, D. Boland, and S. Brewster. 2014. Transient and transitional states: pressure as an auxiliary input modality for bimanual interaction. *Proceedings of the CHI '14,* pp. 401–410. DOI: 10.1145/2556288.2557260. 287, 288

R. McLachlan, M. McGee-Lennon, and S. Brewster. 2012. The sound of musicons: investigating the design of musically derived audio cues. *Proceedings of the ICAD '12*, pp. 148–155. 292

A. Meschtscherjakov, M. Tscheligi, D. Szostak, R. Ratan, R. McCall, I. Politis, and S. Krome. 2015. Experiencing autonomous vehicles: crossing the boundaries between a drive and a ride. *CHI 2015 Extended Abstracts*, pp. 2413–2416. DOI: 10.1145/2702613.2702661. 306

T. Murakami and N. Nakajima. 1994. Direct and intuitive input device for 3-D shape deformation. *Proceedings of the CHI '94*, pp. 465–470. DOI: 10.1145/191666.191823. 287

F. Naujoks, C. Mai, and A. Neukum. 2014. The effect of urgency of take-over requests during highly automated driving under distraction conditions. *Proceedings of the AHFE '14*, pp. 431–438. 306

M. A. Nees and B. N. Walker. 2007. Listener, task and auditory graph: toward a conceptual model of auditory graph comprehension. *Proceedings of the ICAD '07*, pp. 266–273. 293

A. Ng, S. Brewster, and J. H. Williamson. 2013. The impact of encumbrance on mobile interactions. *Proceedings of the INTERACT '13*, pp. 92–109. 302

A. Ng, S. A. Brewster, and J. H. Williamson. 2014. Investigating the effects of encumbrance on one- and two- handed interactions with mobile devices. *Proceedings of the CHI '14*, pp. 1981–1990. DOI: 10.1145/2556288.2557312. 302

J. G. Neuhoff, G. Kramer, N. W. B. Lane, and J. Wayand. 2000. Sonification and the interaction of perceptual dimensions: Can the data get lost in the map? *Proceedings of the ICAD '00*. 294

I. Oakley, M. R. McGee, S. Brewster, and P. Gray. 2000. Putting the feel in "look and feel". *Proceedings of the CHI '00*, pp. 415–422. DOI: 10.1145/332040.332467. 286

S. Pauletto and A. Hunt. 2005. A comparision of audio and visual analysis of complex time-series data sets. *Proceedings of the ICAD '05*, pp. 175–181. 292

B. Plimmer, A. Crossan, S. Brewster, and R. Blagojevic. 2008. Multimodal collaborative handwriting training for visually-impaired people. *Proceedings of the CHI '08*, p. 393. DOI: 10.1145/1357054.1357119. 299, 300

B. Plimmer, P. Reid, R. Blagojevic, A. Crossan, and S. Brewster. 2011. Signing on the tactile line. *ACM Transactions on Computer-Human Interactions,* 18(3):1–29. DOI: 10.1145/1993060.1993067. 299, 300

I. Politis, S. Brewster, and F. Pollick. 2013. Evaluating multimodal driver displays of varying urgency. *Proceedings of the Auomotive UI '13*, pp. 92–99. DOI: 10.1145/2516540.2516543. 296, 297, 305

I. Politis, S. Brewster, and F. Pollick. 2014. Evaluating multimodal driver displays under varying situational urgency. *Proceedings of the CHI '14*, pp. 4067–4076. DOI: 10.1145/2556288.2556988. 305

I. Politis, S. Brewster, and F. Pollick. 2015a. Language-based multimodal displays for the handover of control in autonomous cars. *Proceedings of the Automotive UI '15*, pp. 3–10. DOI: 10.1145/2799250.2799262. 296, 306

I. Politis, S. Brewster, and F. Pollick. 2015b. To beep or not to beep? Comparing abstract versus language-based multimodal driver displays. *Proceedings of the CHI '15*, pp. 3971–3980. DOI: 10.1145/2702123.2702167. 296

A. Potamianos. 2017. Audio and Visual Modality Combination in Speech Processing Applications. In S. Oviatt, B. Schuller, P. Cohen, D. Sonntag, G. Potamianos, and A. Krüger, editors, *Handbook of Multimodal-Multisensor Interfaces, Volume 1: Foundations, User Modeling, and Common Modality Combinations*. Morgan & Claypool, Williston, VT. 288

I. Poupyrev, S. Maruyama, and J. Rekimoto. 2002. Ambient touch: designing tactile interfaces for handheld devices. *Proceedings of the UIST '02*. DOI: 10.1145/571985.571993. 303

C. Rendl, D. Kim, P. Parzer, et al. 2016. FlexCase: enhancing mobile interaction with a flexible sensing and display cover. *Proceedings of the CHI '16*, ACM Press, pp. 5138–5150. DOI: 10.1145/2858036.2858314. 288

S. Robinson, C. Coutrix, J. Pearson, et al. 2016. Emergeables: deformable displays for continuous eyes-free mobile Interaction. *Proceedings of the CHI '16*, ACM Press, pp. 3793–3805. DOI: 10.1145/2858036.2858097. 307

A. Roudaut, A. Karnik, M. Löchtefeld, and S. Subramanian. 2013. Morphees: toward high "shape resolution" in self-actuated flexible mobile devices. *Proceedings of the CHI '13*, pp. 593–602. DOI: 10.1145/2470654.2470738. 307

SAE International Standard J3016. 2014. Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. 306

D. D. Salvucci. 2001. Predicting the effects of in-car interface use on driver performance: an integrated model approach. *International Journal of Human-Computer Studies,* 55(1):85–107. DOI: 10.1145/365024.365064. 304

C. Schwesig, I. Poupyrev, and E. Mori. 2004. Gummi: a bendable computer. *Proceedings of the CHI '04*, pp. 263–270. DOI: 10.1145/985692.985726. 287

P. Sines and B. Das. 1999. Peltier Haptic Interface (PHI) for improved sensation of touch in virtual environments. *Virtual Reality*, 4(4):260–264. DOI: 10.1007/BF01421809. 284

S. D. Speeth. 1961. Seismometer sounds. *The Journal of the Acousticical Society of America*, 33(7):909. DOI: 10.1121/1.1908843. 292

M. A. Srinivasan and C. Basdogan. 1997. Haptics in virtual environments: taxonomy, research status, and challenges. *Computers and Graphics* 21(4):393–404. DOI: 10.1016/S0097-8493(97)00030-7. 286

S. Strachan, P. Eslambolchilar, and R. Murray-Smith. 2005. gpsTunes - controlling navigation via audio feedback. *Proceedings of the Mobile HCI '05*, ACM Press, pp. 275–278. DOI: 10.1145/1085777.1085831. 302

H. Summala, D. Lamble, and M. Laakso. 1998. Driving experience and perception of the lead car's braking when looking at in-car targets. *Accident Analysis Prevention*, 30(4):401–407. DOI: 10.1016/S0001-4575(98)00005-0. 304

K. Tahiroğlu, T. Svedström, V. Wikström, S. Overstall, J. Kildal, and T. Ahmaniemi. 2014. SoundFLEX: designing audio to guide interactions with shape-retaining deformable interfaces. *Proceedings of the ICMI '14*, pp. 267–274. DOI: 10.1145/2663204.2663278. 287

J. B. F. van Erp, K. Kyung, S. Kassner, J. Carter, S. Brewster, G. Weber, and I. Andrew. 2010. Setting the standards for haptic and tactile interactions: ISO's work. *Proceedings of the EuroHaptics 2010 in Lecture Notes in Computer Science 6192 Part 2*, Springer, pp. 353–358. DOI: 10.1007/978-3-642-14075-4_52. 279, 281

J. B. F. van Erp, A. Toet, and J. B. Janssen. 2015. Uni-, bi- and tri-modal warning signals: Effects of temporal parameters and sensory modality on perceived urgency. *Safety Science*, 72:1–8. DOI: 10.1016/j.ssci.2014.07.022. 296

J. B. F. van Erp and H. A. H. C. van Veen. 2001. Vibro-tactile information presentation in automobiles. *Proceedings of the Eurohaptics '01*, pp. 99–104. 305

J. B. F. van Erp, H. A. H. C. van Veen, C. van Jansen, and T. Dobbins. 2005. Waypoint navigation with a vibrotactile waist belt. *ACM Transactions of Applied Perception,* 2(2):106–117. DOI: 10.1145/1060581.1060585. 300

B. N. Walker and G. Kramer. 2005. Mappings and metaphors in auditory displays. *ACM Transactions of Applied Perception,* 2(4):407–412. DOI: 0.1145/1101530.1101534. 293

B. N. Walker, G. Kramer, and D. M. Lane. 2000. Psychophysical scaling of sonification mappings. *Proceedings of the ICAD '00,* pp. 99–104. 294

M. Weiss, S. Voelker, J. Borchers, and C. Wacharamanotham. 2011. FingerFlux: near-surface haptic feedback on tabletops. *Proceedings of the UIST '11*, pp. 615–620. DOI: 10.1145/2047196.2047277. 285

C. D. Wickens. 2002. Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3(2):159–177. DOI: 10.1518/001872008X288394. 278

J. Williamson, R. Murray-Smith, and S. Hughes. 2007. Shoogle: excitatory multimodal interaction on mobile devices. *Proceedings of the CHI '07*, pp. 121–124. DOI: 10.1145/1240624.1240642. 293

G. Wilson, S. Brewster, M. Halvey, and S. Hughes. 2012. Thermal icons: evaluating structured thermal feedback for mobile interaction. *Proceedings of the MobileHCI '12*, pp. 309–312. DOI: 10.1145/2371574.2371621. 279, 284, 285, 297, 631

G. Wilson, S. Brewster, M. Halvey, and S. Hughes. 2013. Thermal feedback identification in a mobile environment. *Proceedings of the HAID 2013*, Article 2. DOI: 10.1007/978-3-642-41068-0_2. 284, 285

G. Wilson, G. Davidson, and S. Brewster. 2015. In the heat of the moment: subjective interpretations of thermal feedback during interaction. *Proceedings of the CHI '15*, pp. 2063–2072. DOI: 10.1145/2702123.2702219. 284, 285

G. Wilson, M. Halvey, S. A. Brewster, and S. A. Hughes. 2011. Some like it hot? Thermal feedback for mobile devices. *Proceedings of the CHI '11*, pp. 2555–2564. DOI: 10.1145/1978942.1979316. 284

L. Yao, R. Niiyama, and J. Ou. 2013. PneUI: pneumatically actuated soft composite materials for shape changing interfaces. *Proceedings of the UIST '13*. DOI: 10.1145/2501988.2502037. 307

W. Yu and S. Brewster. 2002. Comparing two haptic interfaces for multimodal graph rendering. *Proceedings of the HAPTICS '02*, pp. 3–9. DOI: 10.1109/HAPTIC.2002.998934.

W. Yu, K. Kangas, and S. Brewster. 2003. Web-based haptic applications for blind people to create virtual graphs. *Proceedings of the HAPTICS '03*, pp. 318–325. DOI: 10.1109/HAPTIC.2003.1191301. 298

W. Yu, D. Reid, and S. Brewster. 2002. Web-based multimodal graphs for visually impaired people. *Universal Access and Assistive Technology*, pp. 97–108. DOI: 10.1007/s10209-002-0042-6. 298, 299

C. Zhao, K.-Y. Chen, M. T. I. Aumi, S. Patel, and M. S. Reynolds. 2014. SideSwipe: detecting in-air gestures around mobile devices using actual GSM signal. *Proceedings of the UIST '14,* pp. 527–534. DOI: 10.1145/2642918.2647380.