Sonifying the shape of human body motion using motiongrams

ALEXANDER REFSUM JENSENIUS
fourMs Lab, Department of Musicology, University of Oslo

ROLF INGE GODØY
fourMs Lab, Department of Musicology, University of Oslo

ABSTRACT: The paper presents sonomotiongram, a technique for the creation of auditory displays of human body motion based on motiongrams. A motiongram is a visual display of motion, based on frame differencing and reduction of a regular video recording. The resultant motiongram shows the spatial shape of the motion as it unfolds in time, somewhat similar to the way in which spectrograms visualise the shape of (musical) sound. The visual similarity of motiongrams and spectrograms is the conceptual starting point for the sonomotiongram technique, which explores how motiongrams can be turned into sound using "inverse FFT". The paper presents the idea of shape-sonification, gives an overview of the sonomotiongram technique, and discusses sonification examples of both simple and complex human motion.

HOW can the shapes of human body motion be turned into musical sound? This paper will present one approach to understanding more about music and shape through the creation of auditory displays of video recordings of music-related motion. An auditory display uses sound to communicate information (Hermann, Hunt, & Neuhoff, 2011), and, in our case, such auditory displays make it possible to listen to the shapes of the motion in the video recordings. This project started out with the aim of finding alternative ways to analyse what can be summarised as "music-related body motion," the motion that can be observed in both the performance and perception of music (Jensenius, Wanderley, Godøy, & Leman, 2010). Our previous work on this topic has mainly involved the creation of visual displays based on either video recordings or marker-based motion capture data (Jensenius, 2007; Jensenius, 2013). Such visual displays can reveal some of the spatial and temporal content of motion sequences, but usually not all in one display. We have therefore been eager to investigate the possibilities of auditory displays of music-related motion.

There have been several examples of the sonification of three-dimensional motion data recorded with marker-based motion capture systems, such as presented in Quek, Verfaille, and Wanderley (2006), Grond, Hermann, Verfaille, and Wanderley (2010), Höner (2011), and Winters, Savard, Verfaille, and Wanderley (2012). This paper, however, will focus on the sonification of regular video recordings of human motion. It can be argued that video recordings are more limited than motion capture data, with lower frame rate and poorer spatial resolution. On the other hand, a regular video camera is a much cheaper, simpler, less obtrusive, and more accessible recording solution than a motion capture system, and a video camera can also easily be used to record in any type of location and setting. The sonification approach to be presented here, referred to as "sonomotiongram," is based on the sonification of "motiongrams" (see Figure 1 for an overview of the motiongram technique). The conceptual starting point of the project was that motiongrams are visually similar to spectrograms. This similarity made us wonder: what would a motiongram sound like, if it were played back as a spectrogram? The paper starts with an overview of related research. Then motiongrams are introduced, followed by an explanation of how motiongrams can be used to create sound. Finally, some examples of both analytical and interactive applications are presented and discussed.

RELATED WORK

The sonomotiongram method is based on what could be called an "inverse FFT" technique. The idea here is to treat an image as if it were a spectrogram, with frequency information on the Y-axis and time on the X-axis, and use this as the basis for (re)synthesising a sound file. There have been numerous implementations of such an idea over the years, perhaps earliest example being the Pattern Playback ma-

chine built by a group of speech researchers in the late 1940s (Cooper, Liberman, & Borst, 1951). This system made it possible to draw shapes that could afterwards be played back as sound. Iannis Xenakis worked on similar ideas as a compositional strategy, developing the UPIC system (Unité Polyagogique Informatique CEMAMu) in 1977 (Marino, Serra, & Raczinski, 1993). UPIC made it possible to create complex timbres by drawing with a digital pen on a computer screen. During the last decade this idea of making sound from drawings has been available in the Metasynth software, along with the possibility of sonifying any type of image or photo (Wenger, 1998). The ability to draw music more freely has also seen some research interest over the years as digital pen interfaces have become more accessible. Some recent examples include software such as Music Sketcher (Thiebaut, Healey, & Kinns, 2008) and Different Strokes (Zadel & Scavone, 2006).

A similar approach, but with a different starting point, is the synthesis functionality of audio analysis software like AudioSculpt (Bogaards, Röbel, & Rodet, 2004) and SPEAR (Klingbeil, 2005). Both these applications allow for spectral audio analysis, followed by screen-based manipulation of the spectrograms, and resynthesis of the manipulated image into sound. This makes it possible for researchers and composers to edit the timbral content and temporal evolution of sounds in the visual domain.

One important challenge when converting an image into sound is to decide on a mapping between a set of spatial (and timeless) image features and a set of (non-spatial) temporally evolving audio features. In the sonification examples mentioned above, time is created by sequentially running through the X-axis from left to right. A different approach to handling time from images is that of raster scanning (Yeo & Berger, 2006). Here sound is created by scanning through the image pixels line-by-line, starting from the upper-left corner. Yet a different way of handling time is to use a series of moving images as the input to the sonification, as for example the video sonifications by Filimowicz (2010).

Similar sonification strategies have been used in art installations and realtime applications. The SoundView system, for example, allows for the creation of sound by moving a pointer device over an image (van den Doel, 2003). Here the pointer is used as a "tape-head" to produce sound, following the idea of an auditory information-seeking principle (Zhao, Plaisant, Shneiderman, & Duraiswami, 2004). In the augmented reality installation Scrappler users create sonic patterns by putting (physical) objects on a table (Levin, 2006). The location of the object is tracked sequentially and is used to play back a sound whenever the timeline passes the object.

The examples mentioned so far have mainly been based on the sonification of still images or of moving images with a still character. Our main interest, however, is the sonification of moving images of motion. Early experiments in sonifying moving images can be found in the history of film art and early cinema (Bordwell & Thompson, 1997). The Hungarian-British physicist Dennis Gabor developed in 1946 the Kinematical Frequency Converter in view of investigating time-frequency relationships in sound. Based on the technology of optical sound track in cinema, this device enabled a conversion of graphical images to sound by using a flashing light source that would shine through patterns on the film, resulting in a series of images that would activate a photocell that in turn converted these patterns of light into sound (Loy, 2007). In a more artistic vein, the Scottish-Canadian filmmaker Norman McLaren created a series of short films in which he drew with a pen directly on the sound track of a 35mm film strip (Jordan, 1953).

The advent of computers made it possible to use video cameras as sensors, for example by getting information about only parts of the image. In the 1970s Erkki Kurenniemi developed the electronic music instrument Dimi-O for controlling sound synthesis in real time based on video recordings (Ojanen, Suominen, Kallio, & Lassfolk, 2007). David Rokeby created Very Nervous System (VNS) for interactive music and dance in the 1980s, followed by a software version of the system called SoftVNS (Rokeby, 2002). From the turn of the century there has been an increasing number of accessible software solutions for creating sound from live video, including EyesWeb (Camurri, Hashimoto, Ricchetti, Trocca, Suzuki, & Volpe, 2000), Isadora (deLahunta, 2005), and Max/MSP/Jitter (Zicarelli, 1998). Many of these systems use motion detection to control either sound synthesis or samplers in real time. One such example is Pelletier's (2008) direct mapping of motion flow fields to sound. There are also examples of mobile sonification applications, for example the sonification of train journeys based on slit-scanning (Knees, Pohle, & Widmer, 2012).

FROM MOTION TO SOUND

Mapping Musical Features and Shape

Sonomotiongrams belong to the larger domain of *sonification*, a fast growing field of research and development. Sonification is based on the idea that sound may be an efficient transducer of information in various contexts of human life (Hermann, Hunt, & Neuhoff, 2011). In principle, any information, or any sets of data, from such diverse fields as geology, climate research, medical research, or financial markets, can be efficiently transmitted by sonification, meaning that instead of studying large tables and graphs, we can just listen to the data, and provided well-designed schemes for mapping, quickly per-

ceive important information in the data. Sonification can also be done on the fly, making possible interactive scrutiny of various kinds of data, or can be used for feedback in various control tasks, in what is called *interactive sonification*. Such interactive sonification ranges from the classical example of the Geiger counter (making a sparkling sound reflecting the level of radiation) to beeps in modern cars, computers and household devices, to more elaborate schemes of using sound feedback for orientation and navigation.

The underlying principle in all sonifications is that of mapping from some domain (numerical, visual, kinematic, etc.) to the auditory domain. Needless to say, the crucial point will be how this mapping is done. Given the possibility of arbitrary mapping, that is, that any data can in principle be mapped to any sound, the challenge is that of designing mapping schemes that are somehow perceived as meaningful and useful, and importantly, also as aesthetically well-designed. The underlying mental scheme here is what we like to call *shape-cognition*, the belief that human perception and cognition in general is intimately linked with notions of shape and space, as has been suggested by the ideas of image schemata in morphodynamics (Thom, 1983; Godøy, 1997), cognitive linguistics (Johnson, 1987), as well as auditory perception research (Schaeffer, 1966; Godøy, 2006).

As for music and shape-cognition, we have good reason to claim that notions of shape are ubiquitous in Western musical culture. Expressions of shape (words and graphics) are encountered in innumerable contexts of production (shapes of sound-producing body motion and postures), perception (shapes of rhythmic, melodic, timbral, and dynamic features) and theoretical discourse (both notation-based and signal-based such as waveforms and spectrograms). Notions of shape are at the very heart of Western music theory, with the development of notation being clearly shape-oriented already with neumes, however later evolving into more abstract representations. But in spite of this more symbol-oriented evolution of Western music notation, we could say that the performance of a score in a sense is a sonification of the graphics of the score, and, this is an essential point, the graphics of the score are transduced to sound by the musicians' sound-producing body motion.
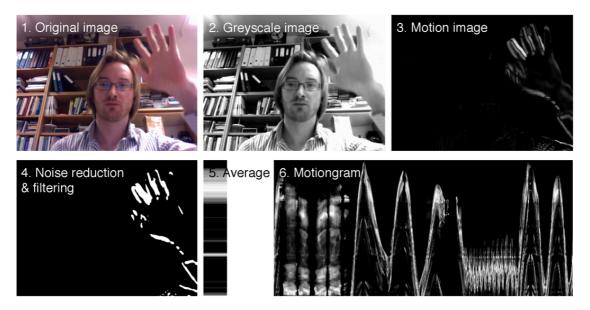
We have in the latter half of the 20th century seen various projects of graphical notation that are less dependent on discrete symbols and rather try to represent continuous trajectories of sound and body motion. At the same time readily available technology has increasingly enabled the study of continuous sound and motion without being limited to a symbol-based system such as the Western music notation. This opens up for study more direct feature-mappings between sound, body motion, and vision in music. The essential element here is that shapes are intrinsically holistic, whereas symbols are not. What we call shape-sonifications may thus capture the ephemeral elements of musical experience, both of sound and music-related body motion, so that we may systematically explore qualitative features and nuances of sound and body motion, and also higher-level affective and aesthetic features.

One important element here is that body motion seems to play a role as translator between modalities. This was (albeit implicitly) suggested by Wolfgang Köhler's famous *Maluma* vs. *Takete* distinction (Köhler, 1947), however a more broad basis for such cross-modal translation has been provided by recent findings in neuroscience suggesting that both visual and sound perception relate closely to motor sensations and simulations (Kohler, Keysers, Umiltà, Fogassi, Gallese, & Rizzolatti, 2002), something that has for several decades been claimed by motor theories of perception (Galantucci, Fowler, & Turvey, 2006). Keeping this in mind, we believe there is the triangular relationship of sound, motion, and vision that can be exploited in the design and use of sonomotiongrams.

Motiongram

A motiongram is a spatiotemporal display of motion in a video recording (Jensenius, 2006; Jensenius, 2013). As the overview in Figure 1 shows, the process starts by reading a video stream and converting it into a greyscale image. In future research it would be interesting to use the colour information as well. It may also be useful to do some simple image adjustments at this stage, such as changing the brightness and contrast, so that the video used for further analysis is as clear as possible (Figure 1.2). The next step involves producing the motion image by calculating the absolute difference between subsequent video frames (Figure 1.3). Dependent on the quality of the original image, and the noise level in the image due to video compression, lighting and so on, it may be necessary to filter the motion image (Figure 1.4). This can be done through simple thresholding or by applying a noise-removal algorithm to remove groups containing few pixels. The motiongram is created by calculating the normalised mean value for each row in the motion image (Figure 1.5). This means that for each image matrix of size *MxN*, a 1*xN* matrix is calculated. Drawing these 1-pixel wide stripes next to each other over time results in a horizontal motiongram (Figure 1.6).

The sonomotiongram technique is based on the concept of treating the motiongram as if it were a spectrogram of an audio recording. Technically, however, a motiongram is very different from a spectrogram. A spectrogram of an audio recording displays the energy level of the frequency bands resulting from doing a Fourier transform on the audio. A motiongram, on the other hand, is a reduced display of a series of motion images. There is no analysis being done when creating a motiongram: It is only based

on a simple reduction algorithm. This reduction approach may be seen as problematic in some contexts. It is, for example, not possible to separate motion in the foreground of the image from motion in the background. On the other hand, the simplicity of the approach also means that it can easily be applied to widely different video material, everything from music and dance performances to the motion of children with cerebral palsy (Jensenius, 2007).

Fig. 1. The steps involved in creating a motiongram: (1) accessing the original video stream, (2) greyscale conversion, (3) frame differencing, (4) filtering, (5) averaging each row, (6) drawing the average matrices over time.

It is worth mentioning that a single motiongram will only display motion in one spatial dimension. Thus a horizontal motiongram visualises only vertical motion, since all information about the spatial distribution of motion in the horizontal plane is represented by 1 pixel for each row. When creating motiongrams it is therefore necessary to evaluate in which plane(s) the motion is occurring, before deciding whether to create a horizontal or a vertical motiongram (or both).

From Motiongram to Sound

Even though motiongrams and spectrograms represent different features, they share one property: the temporal unfolding of shapes of either motion or sound. Furthermore, the Y-axis in a motiongram represents vertical motion, which is often associated with pitch. We were therefore curious to hear what would happen if it were treated as a spectrogram and turned into sound, as illustrated in Figure 2.
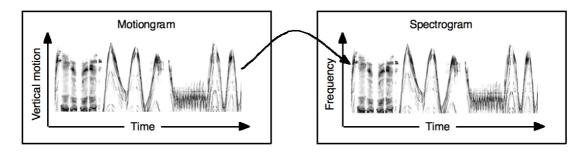

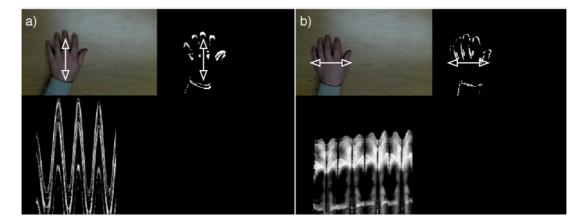**Fig. 2.** A sketch of the direct mapping from motiongram to spectral audio data.

The sonomotiongram technique has been implemented in the graphical programming environment Max/MSP/Jitter, and the technical details have been presented in Jensenius (2012). The implementation is based on passing data from the motiongram matrix to an interpolated oscillator bank performing the additive sound synthesis. The result is a direct sonification of the motion, where lower sound frequencies are controlled by movement in the lower part of the image, and vice versa.

The sonification algorithm was implemented in the module jmod.sonifyer~ of the open framework Jamoma for Max (Place & Lossius, 2006). The modular approach of Jamoma makes it possible to create complex audio and video patches in Max quickly, with the added benefit of extensive preset, mapping and cueing functionality (Place, Lossius, Jensenius, & Peters, 2008). An example of how the sonifyer module may be used in conjunction with other Jamoma modules is shown in Figure 3, and a video tutorial of the functionality of the module can be seen in Video 1.

Fig. 3. An example of how a set of Jamoma modules can be used to sonify a video stream in realtime. The video input module in the upper left corner is connected to a module creating the motion image and then to the motiongram module. Finally, the output motiongram is sent to the sonifyer module for the creation of sound.

Sonification of Basic Motion Patterns

Before presenting some examples of the sonification of complex motion, it may be useful to investigate some basic motion patterns. Video 2 shows examples of moving the hand up and down, sideways, diagonally and in a circle. These examples effectively display the largest problem with the proposed sonification approach: the motiongram's ability to display motion in only one direction. Thus the up- and downward motion is clearly visualised in the motiongram (Figure 4), and heard in the sound. The other motion patterns (sideways, diagonal and circular) are not represented equally well because the motion is only partly happening in the horizontal direction, is therefore not properly visualised in the motiongram, and hence is not audible.
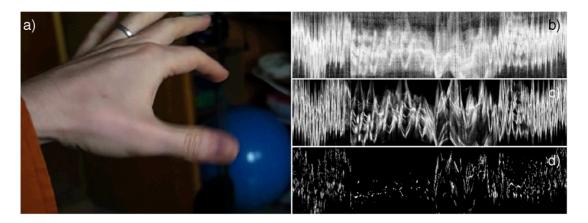

Fig. 4. Screenshots from Video 2 showing: (a) the motion pattern of a hand moving up and down, (b) the motion pattern of a hand moving sideways. The arrows are added manually to indicate the direction of the motion. Notice how the motiongram in (a) effectively visualises the motion, while the motiongram in (b) only visualises the (small) vertical displacement of the motion.

One attempt at sonifying the two motion axes at the same time is shown in Video 3. Here both horizontal and vertical motiongrams are created from the same video recording, and the sonifications of the two motiongrams have been mapped to the left and right audio channel respectively. While this method gives some impression of the motion along the two axes, understanding the relationship be-

tween the two may be quite confusing. Clearly, more experimentation is needed to find a better solution for sonifying the two dimensions independently yet simultaneously.

Effect of Filtering

One of the most important decisions to make when creating a motiongram, and hence its sonification, is the level of filtering and thresholding to apply to the image. Setting a high threshold value will remove more noise from the image but it may also remove valuable parts of the motion to be sonified. This can be seen in an example of a high-speed recording (200 fps) of a hand in motion in Video 4 and Figure 5. Here three different types of filtering have been applied to the motion image to show the visual affect on the motiongram and how it influences the sonic results. When there is no thresholding and no noise re-duction, all the details of the hand motion are visible and audible. Adding a low-pass filter removes a substantial quantity of pixels in the motiongram, and hence makes a cleaner sonification. Finally, adding a noise reduction algorithm further reduces the number of pixels, leading to a sonification of only the most important parts of the motion sequence.

Fig. 5. Examples of how filtering the motion image influences the motiongram: (a) screenshot from the original video recording, (b) no filtering, no noise reduction, (c) low-pass filtering, (d) low-pass filtering and noise reduction.
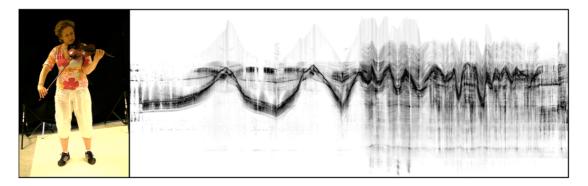
EXAMPLES OF REALTIME SONIFICATION

Let us now compare sonification of different types of music-related body motion. The examples in this section will focus on realtime sonifications, that is, the sound of the sonification is temporally correlated with the performer's motion.

Sonification of violin performance

Sonifying the motion of a musician may seem like a paradox, since the sound-producing and sound-modifying actions of a performer are in themselves sound-producing. In these examples, however, we will look at the overall body motion of a performer, not only the sound-producing actions. Video 5 and Figure 6 show an example of the sonification of a short violin improvisation going from slow legato bowing to more rapid and energetic strokes. One confusing element here is that the first part of the soni-fication resembles the original sound of the performance, since it is mainly the bowing arm that is soni-fied in the beginning of the sequence. In the second half of the sonification, however, motion in the rest of the body is also sonified. As such, sonifying musicians' motion with such a generic sonification tech-nique may lead to results that are somewhat tricky to interpret, since both sound-producing and non-sound-producing motion are sonified. But this may also be seen as one of the strengths of the technique, namely the ability to put into sound all the visible elements of a performance.

Sonification of Clogging

While the sonomotiongram technique is not based on any particular type of analysis or musical knowl-edge, it is possible to improve the sonification results manually by adjusting the visual input to the sys-tem. An example of this can be seen in the sonification of the motion of a French-Canadian fiddler in Video 6 and Figure 7. Instead of sonifying the full body motion of the performer, we are here focusing on the clogging pattern of his feet. This makes it possible to listen to, for example, the rhythmic struc-

ture of the clogging pattern, and pick up changes such as the change of rhythmic figure and tempo half-way throughout the excerpt. This change is not easy to see in the original video, nor in the motiongram, but is clearly audible in the sonification. (See Schoonderwaldt & Jensenius (2011) for a rationale and a more detailed analysis of this performance.)

Fig. 6. Sonification of a short violin performance, starting with slow legato strokes and moving to more energetic strokes in the second half of the improvisation (43 seconds).
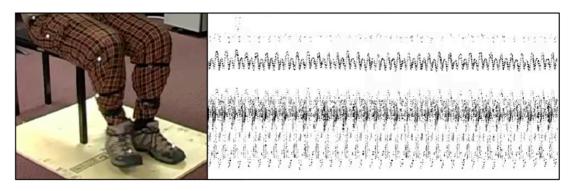

Fig. 7. Sonification of the clogging pattern of a French-Canadian fiddler (30 seconds).

Sonification of Dance Motion

An example of the sonification of dance motion is shown in Figure 8. This recording is from an experiment in which a group of dancers were asked to improvise freely to short excerpts of music. Video 7 starts by showing the original recording of one of the dancers, and how the dancer moves spontaneously to the sound of music. Then follows the same video but overlaid with the sonification of her motion. It is interesting to hear how the sonification of the motion shows some clear similarities to the sonic qualities of the original sound. This, however, is a special case of a good correspondence between the original sound and the sonification result. Although such correspondences may happen, we find that it is often more interesting to use sonification as a method to reveal motion features that contrast with the original audio, or to reveal motion features that are not easily visible to the naked eye. In a non-realtime version of our software, such parts of a sonomotiongram may be manually investigated by scrubbing through the motiongram with the mouse.
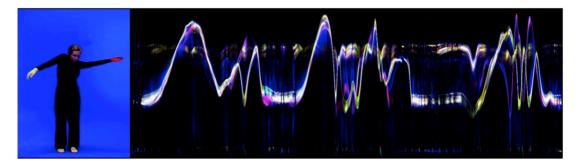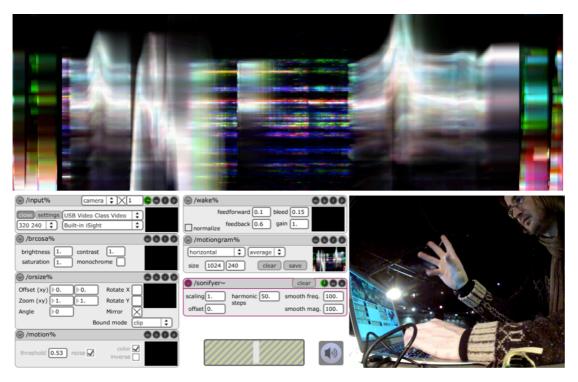

Fig. 8. Sonification of free dance motion (64 seconds).

Sonification in Live Performance

Sonification of body motion is usually seen as something different than musical mapping of the same body motion (Winters, Savard, Verfaille, & Wanderley, 2012). That is, the aim of a sonification is to convey information about the data, not to create musically and aesthetically interesting sound. But that does not close the door to using a sonification, or a sonification technique, in a musical context. While the sonomotiongram technique was originally developed for analytic applications, we have found that its greatest potential may, in fact, be for creative applications. Since the Jamoma sonification module runs in real time, it can easily be included in a performance setup. Over the last couple of years the module has been used in several concerts, both in solo performance (Video 8) and with ensembles (Video 9). Figure 9 shows the performance patch for the piece *Soniperforma*, which uses the sonification module as the main sound maker. The processing is based on applying various video effects to the input video, such as changing the colours, size and orientation of the image, and applying image filters and video feedback. Since the sonification is based on the input image, such video effects will in practice end up as sound effects. For example, adding a motion blur effect to the video image will result in a delay in the sound. We have found this to be creatively inspiring as a performer, and also engaging for audiences.

Fig. 9. A performance patch using various video effects to modify the motiongram, and hence the output sound. The visual result from changing various video parameters can be seen in the motiongram at the top.

CONCLUSIONS

Given the basic tenet that very many (or most) musical features can be represented as shapes, we have the basis for systematic explorations of perceptually salient musical features by shape-sonification. Available technologies for capturing visual shapes (video recordings and motion capture), for image processing/synthesis (still pictures, animations, robotics), and for sound synthesis, make it possible to generate incrementally different variants of sonifications. Such sonifications can be informative in themselves, and can also be evaluated and used in experiments on similarity judgments. In other words, with shape-sonification we can embark on a systematic *analysis-by-synthesis* approach to music and shape. The sonomotiongram technique is one approach to creating such shape-sonifications.

In this work, we may assume that there is a continuum from more direct physical sound-motion-shape relationships (signal-based), by way of various corporeal sound-motion-shape relationships (both sound-producing and sound-accompanying body motion) to more high-level, aesthetic and affective music-shape relationships. This should also make possible investigations of consensus on

music-shape relationships among larger groups of perceivers; it is for instance not evident that there should be a cross-cultural consensus on all music-shape relationships, and this is one of the issues that we wish to look into in the future. One important factor here is that cross-modal correspondences (or resemblances) may be approximate. In our own work on music-related body motion, we have seen that there may be a clear consensus on certain music-motion shape relationships; however there may be variations in detail. Yet in assessing such degrees of consensus versus disagreements, we believe also approximate similarity is important and probably also intrinsic to musical experience, in the way a visual sketch may very well reflect the gist of a scene or personality in spite of its rather sparse content.

Systematic investigations of music-shape relationships by sonifications engage in a kind of "hermeneutic circle" with shifts between graphical renderings of musical features as shapes, and conversely, sonifications of graphical shapes. We expect that such systematic shifts between relatively solid visual images and ephemeral sound and body motion can tell us more about salient musical features and about music as a phenomenon.

END MATTER

Acknowledgments

REFERENCES

Bogaards, N., Röbel, A., & Rodet, X. (2004). Sound analysis and processing with Audiosculpt 2. In: *Proceedings of the International Computer Music Conference*. Miami, Florida, pp. 462-465.

Bordwell, D., & Thompson, K. (1997). *Film Art. An Introduction*. New York: McGraw-Hill Companies, Inc.

Camurri, A., Hashimoto, S., Ricchetti, M., Trocca, R., Suzuki, K., & Volpe, G. (2000). Eyesweb: Toward gesture and affect recognition in interactive dance and music systems. *Computer Music Journal*, Vol. 24, No. 1, pp. 57-69.

Cooper, F., Liberman, A., & Borst, J. (1951). The interconversion of audible and visible patterns as a basis for research in the perception of speech. In: *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 37, No. 5, pp. 318.

deLahunta, S. (2005). Isadora almost out of beta: tracing the development of a new software tool for performing artists. *International Journal of Performance Arts and Digital Media*, Vol. 1, No. 1, pp. 31-46.

Filimowicz, M. (2010). Video sonification. http://www.filimowicz.com/pjim/.

Galantucci, B., Fowler, C.A., & Turvey, M.T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin and Review*, Vol. 13, No. 3, pp. 361-377.

Godøy, R.I. (1997). *Formalization and Epistemology*. Oslo: Scandinavian University Press.

Godøy, R.I. (2006). Gestural-sonorous objects: embodied extensions of Schaeffer's conceptual apparatus. *Organised Sound*, Vol. 11, No. 2, pp. 149-157.

Grond, F., Hermann, T., Verfaille, V., & Wanderley, M.M. (2010). Methods for effective sonification of clarinetists' ancillary gestures. *Gesture in Embodied Communication and Human-Computer Interaction*, pp. 171-181.

Hermann, T., Hunt, A., & Neuhoff, J.G. (2011). *The Sonification Handbook*. Berlin: Logos Verlag.

Höner, O. (2011). Aiding movement with sonification in "exercise, play and sport". In: T. Hermann, A. Hunt, & J.G. Neuhoff, (Eds.), *The Sonification Handbook*. Berlin: Logos Verlag, pp. 525-553.

Jensenius, A.R. (2006). Using motiongrams in the study of musical gestures. In: *Proceedings of the International Computer Music Conference*, New Orleans, LA, pp. 499-502.

Jensenius, A.R. (2007). *Action–Sound: Developing Methods and Tools to Study Music-Related Body Movement*. PhD dissertation. University of Oslo.

Jensenius, A.R. (2012). Motion-sound interaction using sonification based on motiongrams. In: *Proceedings of the Fifth International Conference on Advances in Computer-Human Interactions*, Valencia, pp. 170-175.

Jensenius, A.R. (2013). Some video abstraction techniques for displaying body movement in analysis and performance. *Leonardo*, Vol. 46, No. 1, pp. 53-60.

Jensenius, A.R., Wanderley, M.M., Godøy, R.I., & Leman, M. (2010). Musical gestures: Concepts and methods in research. In: R.I. Godøy & M. Leman, (Eds.), *Musical gestures: Sound, movement, and meaning*, New York: Routledge, pp. 12-35.

Johnson, M. (1987). *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.

Jordan, W.E. (1953). Norman McLaren: His career and techniques. *The Quarterly of Film Radio and Television*, Vol. 8, No. 1, pp. 1-14.

Klingbeil, M. (2005). Software for spectral analysis, editing, and synthesis. In: *Proceedings of the International Computer Music Conference*, Barcelona, pp. 107-110.

Knees, P., Pohle, T., & Widmer, G. (2012). Sound/tracks: artistic real-time sonification of train journeys. *Journal on Multimodal User Interfaces*, Vol. 6, No. 1, pp. 87-93.

Kohler, E., Keysers, C., Umiltà, M.A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, No. 297, pp. 846-8.

Köhler, W. (1947). *Gestalt Psychology, 2nd Ed.*. New York: Liveright.

Loy, G. (2007). *Musimathics: the mathematical foundations of music, Volume 2*. Cambridge, MA: MIT Press.

Levin, G. (2006). The table is the score: An augmented-reality interface for real-time, tangible, spectrographic performance. In: *Proceedings of the International Computer Music Conference*, New Orleans, LA, pp. 151-154.

Marino, G., Serra, M.-H., & Raczinski, J.-M. (1993). The upic system: Origins and innovations. *Perspectives of New Music*, Vol. 31, No. 1, pp. 258-269.

Ojanen, M., Suominen, J., Kallio, T., & Lassfolk, K. (2007). Design principles and user interfaces of erkki kurenniemi's electronic musical instruments of the 1960's and 1970's. In: *Proceedings of the International Conference on New Interfaces for Musical Expression*, New York, pp. 88-93.

Pelletier, J.-M. (2008). Sonified motion flow fields as a means of musical expression. In: *Proceedings of the International Conference on New Interfaces For Musical Expression*, Genoa, pp. 158-163.

Place, T., & Lossius, T. (2006). Jamoma: A modular standard for structuring patches in Max. In: *Proceedings of the International Computer Music Conference*, New Orleans, LA, pp. 143-146.

Place, T., Lossius, T., Jensenius, A.R., & Peters, N. (2008). Flexible control of composite parameters in max/msp. *Proceedings of the International Computer Music Conference*, Belfast, pp. 233-236.

Quek, O., Verfaille, V., & Wanderley, M.M. (2006). Sonification of musician's ancillary gestures. *Proceedings of the International Conference on Auditory Display*, London, pp. 194-197.

Rokeby, D. (2002). softVNS [software]. http://homepage.mac.com/davidrokeby/softVNS.html.

Schaeffer, P. (1966). *Traité des objets musicaux*. Paris: Éditions du Seuil.

Schoonderwaldt, E. & Jensenius, A.R. (2011). Effective and expressive movements in a French-Canadian fiddler's performance. In: *Proceedings of the International Conference on New Interfaces for Musical Expression*, Oslo, pp. 256-259.

Thiebaut, J.-B., Healey, P.G.T., & Kinns, N.B. (2008). Drawing electroacoustic music. In: *Proceedings of the International Computer Music Conference*, Belfast, pp. 452-458.

Thom, R. (1983). *Paraboles et catastrophes*. Paris: Flammarion.

van den Doel, K. (2003). Soundview: Sensing color images by kinesthetic audio. In: *Proceedings of the International Conference on Auditory Display*, Boston, MA, pp. 303-306.

Wenger, E. (1998). Metasynth [software]. http://www.uisoftware.com/metasynth/.

Winters, R.M., Savard, A., Verfaille, V., & Wanderley, M.M. (2012). A sonification tool for the analysis of large databases of expressive gesture. *The International Journal of Multimedia & Its Applications*, Vol. 4, No. 6, pp. 13-26.

Yeo, W.S., & Berger, J. (2006). Application of raster scanning method to image sonification, sound visualization, sound analysis and synthesis. In: *Proceedings of the International Conference on Digital Audio Effects*, Montreal, pp. 309-314.

Zadel, M., & Scavone, G. (2006). Different strokes: a prototype software system for laptop performance and improvisation. In: *Proceedings of the International Conference on New Interfaces for Musical Expression*, Paris, pp. 168-171.

Zhao, H., Plaisant, C., Shneiderman, B., & Duraiswami, R. (2004). Sonification of geo-referenced data for auditory information seeking: Design principle and pilot study. In: *Proceedings of the International Conference on Auditory Display*, Sydney, pp. 33-36.

Zicarelli, D. (1998). An extensible real-time signal processing environment for max. In: *Proceedings of the International Computer Music Conference*, Beijing, pp. 463-466.