RASTERPIECE: A CROSS-MODAL FRAMEWORK FOR REAL-TIME IMAGE SONIFICATION, SOUND SYNTHESIS, AND MULTIMEDIA ART

Woon Seung Yeo and Jonathan Berger
Center for Computer Research in Music and Acoustics
Stanford University
Stanford, CA 94305, USA
woony@ccrma.stanford.edu

ABSTRACT

This paper introduces *Rasterpiece*, a multimedia environment for direct real-time sound synthesis from images based on the method of raster scanning combining scanning variables (such as sample rate and probe size) with image processing filters. Rasterpiece enables cross-modal integration of image and sound. By providing an arbitrary number of image windows, each with independent control of sonification parameters and image processing, Rasterpiece supports simultaneous playback, mixing, and recording of multiple sonification results. Applications include diagnostic exploration of visual data, creative sound design and composition, and real-time multimedia performance. Special attention is paid to its use as a sound particle design tool for pulsar synthesis.

1. INTRODUCTION

Raster scanning is a technique for generating or recording the elements of a display image by sweeping the screen in a line-by-line manner. More specifically, it scans the whole area, typically progressing in a top-down direction moving from left to right (figure 1). This technique provides an intuitive and straightforward scheme for mapping between one- and two-dimensional data spaces. Its simple, one-to-one correspondence also makes itself a totally reversible process: data converted into one representation can be reconstructed losslessly .

In earlier papers [22] [23] we proposed raster scanning as a geometric mapping framework for image sonification and sound visualization. This method produces coherent and compelling results when translating a visual texture to an auditory display and, conversely, when representing timbre visually. With its potential as a cross modal representation, the complementary and analogous property of raster mapping can be applied to construct a creative and useful method of audio and image processing.

In this paper we describe *Rasterpiece*, an image sonification software environment that extends the mapping framework of raster scanning both in the audio and visual domains. Rasterpiece allows for selective real-time raster-scanned sonification by mouse-based dragging across an arbitrary portion of an image surface. The software also offers a robust platform for powerful image processing. Currently nine visual filters extend the cross-modality of raster mapping by allowing the user to explore mappings between auditory and visual signal processing. The software integrates audio and visual information with an "image interface" highly evocative of its sonified result. Moreover, by providing an arbitrary number of image windows with independent audio and visual engine, Rasterpiece allows for complexity through visual multiplicity and polyphony.

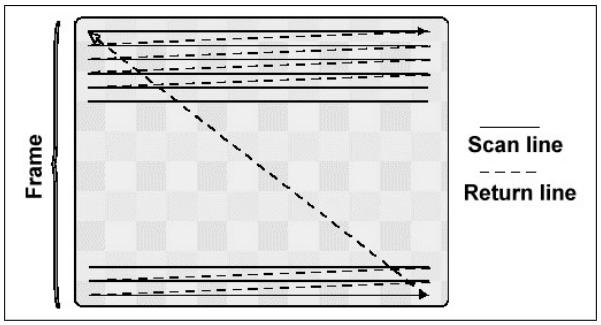


Figure 1. Geometric framework of raster scanning.

The multimedia nature of Rasterpiece suggests a broad range of applications including audification, sonification and visualization for data representation and analysis, as well as a creative tool for multimedia art, a novel means of sound synthesis and a tool for music composition and performance.

1.1. Raster sonification: a brief review

The basic rules of raster mapping for sonification is defined as follows:

- Brightness values of image pixels, (in the case of grayscale images ranging from 0.0 (black) to 1.0 (white)), are linearly scaled to fit into the range of audio sample values from -1.0 to 1.0.
- One image pixel corresponds to one audio sample such that the area of an image corresponds to the duration of its sonified sound.

Figure 2 illustrates these rules.

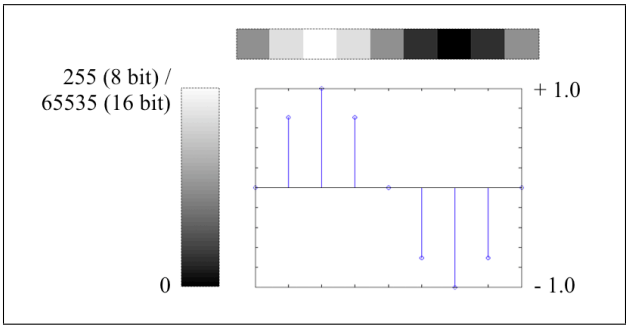


Figure 2. Rules of raster mapping.

When raster-scanning an image with a high degree of pattern periodicity (typical of most visual images) the sonified result will correspondingly have a high degree of perceived pitch wherein the width of an image determines the period (thereby the pitch) of its sonified sound. In general, the width of an image with considerable pattern repetition is inversely proportional to the pitch of its raster-sonified sound. When there is little coherent pixel variation along a horizontal plane (i.e., white noise) the percept of noise overrides that of audible pitch. Additionally, image size is an important factor in raster-scan sonification particularly in the domain of pitch perception. Since the duration of the sonified sound is determined by the image area, small images may produce sounds that are shorter than the threshold of pitch perception [1] . This will be further discussed in section 4.3 with the potential of Rasterpiece as a sound generation tool for pulsar synthesis.

1.2. Mapping between image texture and timbre

Raster mapping provides a method of directly translating between visual texture and timbre in a meaningful way. Figure 3 illustrates four images with contrasting textures. Informal preliminary experiments suggest that when all four sonified mappings are played sequentially in random order listeners generally correctly match each individual image with its respective sonification. Further and more formal exploration and experimentation is planned in hopes of defining a consistent mapping between image and sound texture.

An extension of this mapping involves exploring the effects of raster-based sonification of applying and manipulating visual filters on a given image. Figure 4 contains an image together with selected filtered results. Again, the sonified results of these image variations seem to be consistently perceived as an appropriate auditory display of the visual transformation. From the original image of figure 4(a) the sound of figure 4(b) produces the auditory image described by listeners as 'bumpy', while that of figure 4(c) is described as 'noisy' and 'grainy'. In addition, figure 4(d) results in a sound with modified dynamic range (e.g., by compressor/expander.)

While further investigation of texture-to-timbre mapping is warranted, Rasterpiece has proven to be a novel framework for creative use of visual filters as a means of sound design and timbre control.

[1] Typically 10∼50 [ms], depending on the frequency of sound [13].

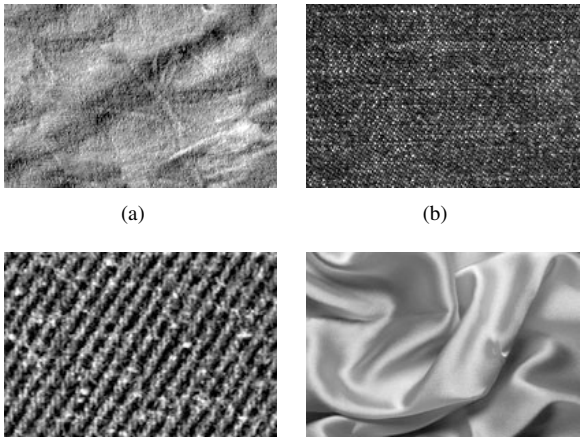


Figure 3. Images used to discriminate different image textures in the auditory domain.

2. RASTERPIECE IN THE CONTEXT OF RELATED ARTISTIC WORK

Numerous approaches to auditory display of image data have been creatively applied by composers and artists [3]. Of these, two general approaches - graphical interface and data mapping - deserve consideration as precedents of sonification using Rasterpiece.

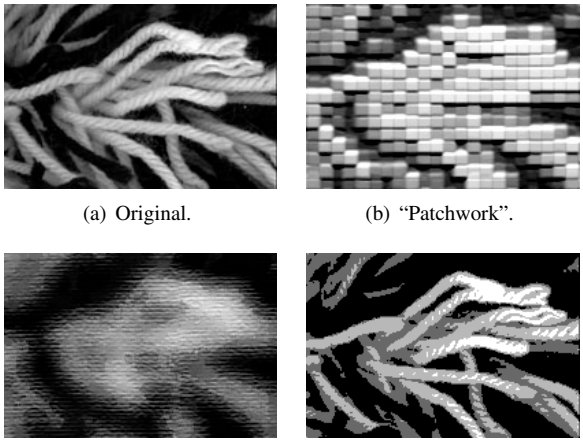
2.1. Graphic synthesis

The mechanisms of translating visual images into sound have been termed *graphic synthesis* [11] [13] . An example of graphic synthesis is *photosonic synthesis* [2], which incorporates a rotating optical disc with its inscribed time-domain waveform: in this approach, a short waveform repeats many times per second to form a tone. Another prime work in the early days is the technique of animated sound developed by McLaren [10]: he "drew" lines and curves on the audio portion of his films to create the sound for the motion pictures.

It should be noted that most well-known examples of graphic synthesis, such as the UPIC system [18], Metasynth [15], Audiosculpt [8], vOICe method [9], and Sonos [14], take a different approach based on frequency-domain representation, often powered by Fourier analysis/synthesis: graphical interface displays the sound in a "time-frequency" plane and allows the user to edit the spectra of sound. This mechanism, sometimes called as *sonographic synthesis*, differs from the simple mapping of raster sonification that does not require any spectral transformation.

2.2. Data mapping by scanning method

Raster sonification is based on a mapping from two- to one-dimensional data space. This is comparable to *wave*



(a) Original. (b) "Patchwork". (c) "Grain". (d) "Posterize".

Figure 4. Comparison of visually filtered images.

terrain synthesis [4], which is based upon a wave surface scanned in an orbit (a closed path), where movement of the orbit causes variations in the generated sound. Although several scanning techniques, including [5], have been explored, they are distinct from raster scanning.

The fact that data is sonified along a pre-defined linear path is reminiscent of the *scanned synthesis* [17], a sound synthesis technique which scans a closed path in a data space periodically to create a sound. Due to its emphasis on the performers control of timbre, data to be scanned is usually generated by a slow dynamic system whose frequencies of vibration are below about 15 [Hz], whereas the pitch is determined by the speed of the scanning function. The system features dynamic wavetable control and is characterized by various scanning patterns, whereas raster sonification features a fixed geometric framework dedicated to converting rectangular images to sound.

2.3. Scanning vs. probing

In [21], we proposed the concept of pointer - path pair and scanning - probing methods as the basis of a general framework for mapping classification. Although raster mapping is an example of scanning along a pre-defined path, Rasterpiece allows for arbitrary free probing of images by supporting raster sonification of an arbitrarily small area. Section 4.1 describes this in detail.

3. FEATURES OF RASTERPIECE

Rasterpiece not only provides a robust framework for image sonification based on the aforementioned raster mapping rules, but also offers additional functionalities for interactive sound synthesis and image processing. Significant features of the software include the following:

- Fast and convenient image sonification. Any rectangular region of an image can be instantly sonified.
- Color information support. Brightness to audio level mapping of Rasterpiece is not restricted to grayscale-only mapping, but offers several color mapping options.
- Playback speed control. Sonified sound can be played at any rate between $f_s/4 \sim 4f_s$, where f_s is the sampling rate of the audio device.
- Real-time image processing. Visual filters are used to modify sound in cross-modal context.
- Simultaneous sonification of multiple images. This is equivalent to playing multi-track audio samples.

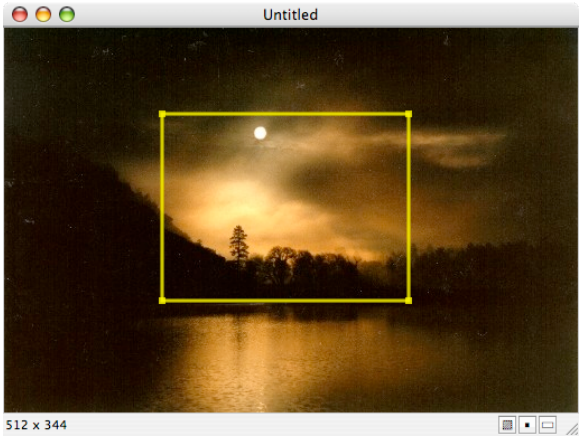


Figure 5. Screenshot of a *canvas*.

Through the rest of this section we describe major components of Rasterpiece and their significance in detail.

3.1. *Canvas* and *marquee*

Figure 5 shows a *canvas* - an image window, which is the most fundamental interface of Rasterpiece.

Inside each canvas lies a *marquee* (yellow rectangle with knobs on four corners) to show currently selected region to be sonified. Like its namesakes in many image processing applications (e.g., Adobe Photoshop, The GIMP, etc.), position and size of a marquee can be freely changed instantly. Furthermore, this selected area can then be sonified in real-time. Sonified sound may be either played just once, or looped for an infinite duration, while being recorded to a sound file simultaneously.

In addition to marquee, buttons to control image, sonification, and marquee settings are provided in the bottom of the window.

3.2. Multiple canvases: simultaneous sonification

Virtually unlimited number of canvases, each with independent control of color mapping and audio settings, can be simultaneously open and active for sonification. A screenshot of multiple image windows is shown in figure 6.

Flexible audio processing framework of Rasterpiece enables the user to play, mix, and record sonified results.

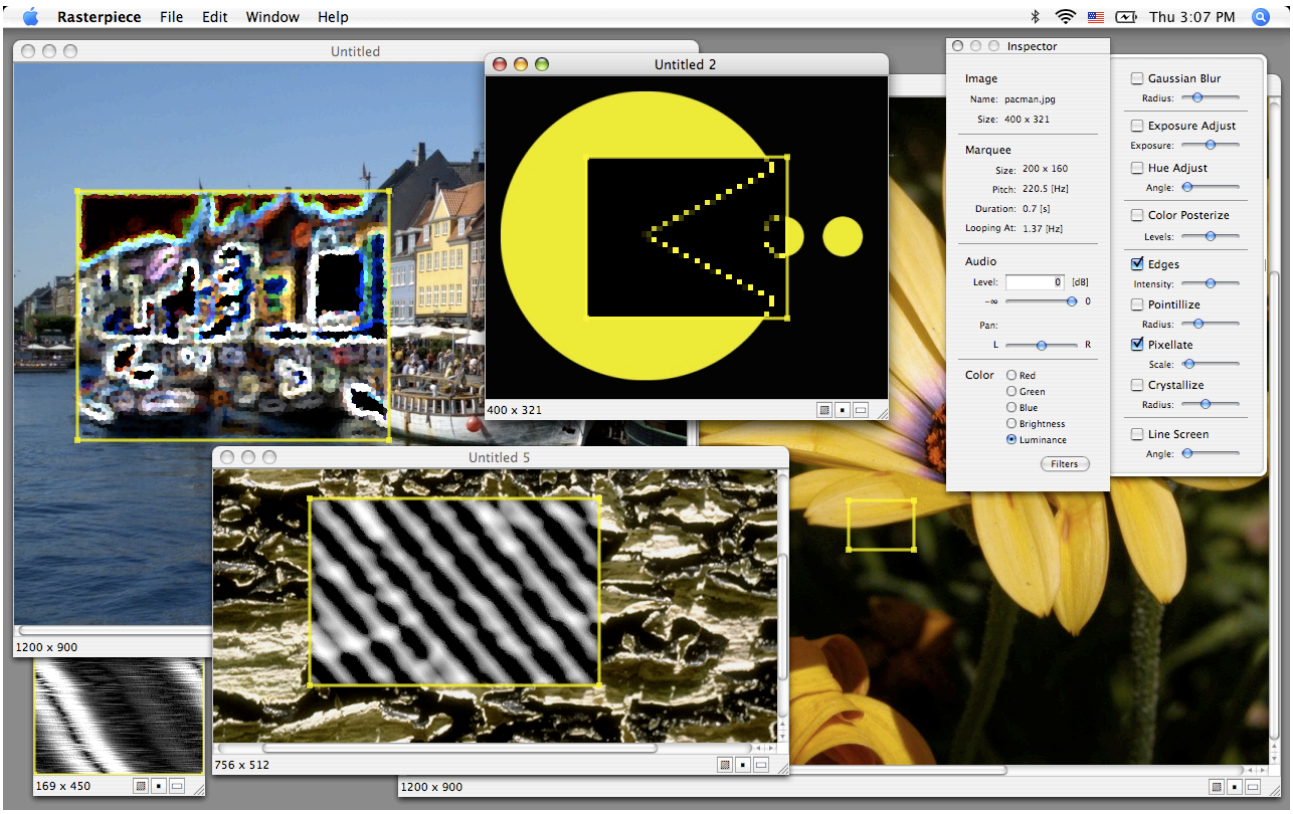


Figure 6. Screenshot of Rasterpiece, with multiple canvases opened simultaneously.

3.3. Inspector: parameter control interface

At the center of the user interface of Rasterpiece is the *inspector*, which displays information on currently selected "active" canvas and offers controls of image processing, data mapping, and audio settings. Most parameters can be monitored and manipulated via this panel (see figure 7).

Inspector consists of the following sections:

- Image (a). The file name and size of the image is displayed.
- Marquee (b). Based on the current marquee size, inspector shows the "nominal" pitch (frequency calculated from the width) and the duration of sonified sound: "looping frequency" corresponds to the inverse of duration. Information provided in this section proves to be useful when "playing" sound by moving/resizing marquee.
- Audio (c). Gain, pan, and playback speed of the sonified result can be controlled here.
- Color (d). In addition to the simple amplitude mapping of grayscale images, one of the RGB components can be chosen as the *color* information of each pixel to be mapped to audio sample value. Also available are *value* (the largest one of RGB), and *luminance* (the representation of brightness that corresponds best to human perception [2] .)

[2] Luminance L of a pixel in RGB format is calculated by $L = 0.30R + 0.59G + 0.11B$.

3.4. Visual filters: real-time image and audio processing

As discussed in section 1.2, raster mapping produces interesting and impressive results when used for sonifying the effects of various visual filters: raster-sonified sound of a visually filtered image is reminiscent of the filter characteristics in the auditory domain. Rasterpiece further extends this idea to the concept of "sound manipulation using image processing techniques," and provides a set of visual filters for both image and sound processing in real-time (see the *filter drawer* - (e) of figure 7). Filters supported by Rasterpiece may be categorized into the followings:

- Focus: Gaussian blur.
- Color adjustment: exposure adjust and hue adjust.
- Color effect: color posterize.
- Stylize: edges, pointillize, pixellate, and crystallize.
- Halftone effects: line screen.

It is well understood that some of these filters may correspond to their audio counterparts. For example, a *blur* filter shows similar characteristics of lowpass audio filters, whereas edge detection can be compared to highpass filtering. In addition, exposure and color adjustments can be understood as dynamic range processes such as normalization and compression. On the contrary, most stylizing

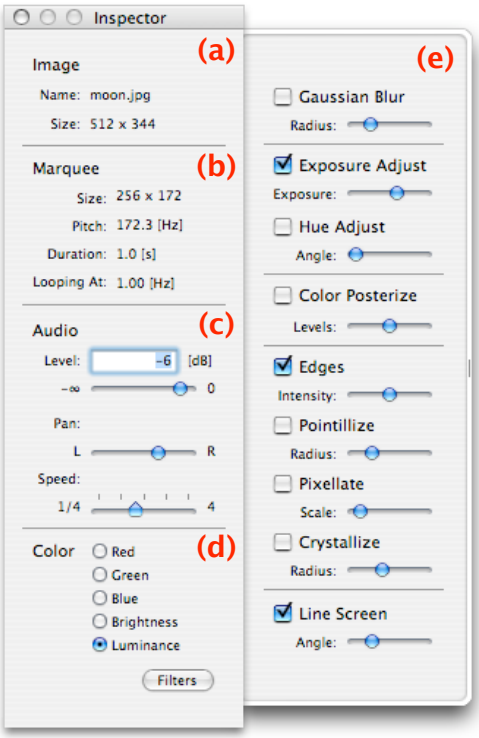


Figure 7. Inspector panel and filter drawer.

filters do not fall onto any specific category of audio filter: typical algorithms for "texturizing" an image are designed to work in a two-dimensional data space, which is rare in the audio domain. In fact, this clearly exemplifies one of the significant features of Rasterpiece: a new creative paradigm for image sonification and sound synthesis in the context of cross-modal mapping. Figure 8 illustrates examples of image filtered marquees. It should be noted that more than one filter may be applied to a marquee at the same time.

The image processing capability of Rasterpiece strongly suggests its potential not only as an image sonification framework, but also as a multimedia environment for cross-modal exploration of combined audio and visual information.

3.5. Notes on software implementation

Rasterpiece is currently implemented as a Cocoa-based Mac OS X application. Its skeletal structure is mostly inherited from *SonART* [20], a network-based sonification framework for layered images and multi-dimensional datasets in a pixel-by-pixel manner.

For image processing, Rasterpiece incorporates built-in image filters supplied by Apple's *Quartz* and *Core Image* graphics framework [1]. This library provides not only powerful visual tools but also easy-to-use interface for developers.

Sound engine of Rasterpiece is powered by the *Synthesis Toolkit* (STK) [6] via *StkX* - an STK framework for OS X [24]. A new class *RstrWvIn* has been written to read bitmap buffer data and convert them into audio samples.

Figure 9. Data probing with Rasterpiece.

4. APPLICATION AREAS

In addition to being a basic experiment tool for the theory of raster sonification, Rasterpiece has versatile use for sound synthesis and performance as well as auditory display of visual information.

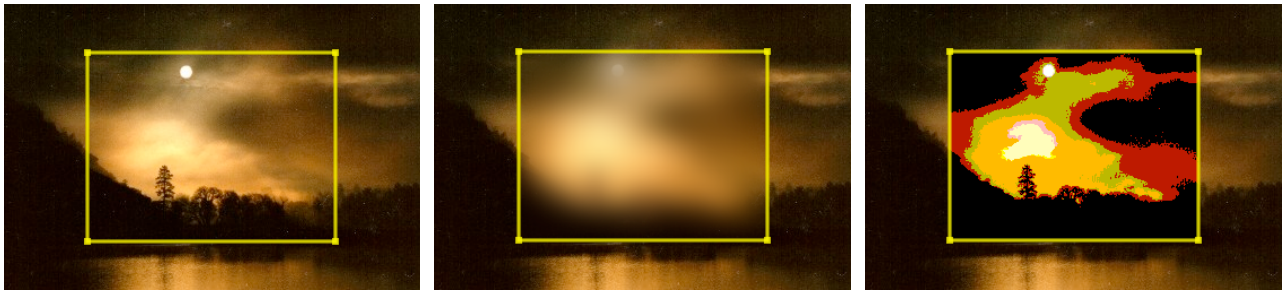
4.1. Diagnostic exploration of visual data

We have mentioned that the mapping mechanism of raster sonification *scans* the whole area by sweeping it along a fixed, pre-determined path. This mapping proves to be highly effective in sonifying the visual pattern and texture of images on a relatively large scale. Certain applications, however, require that "local" characteristics of very small areas of an image need to be examined and compared with others. Possible examples include diagnostics of medical imaging data, in which visually salient areas are *probed* to detect any malignant region with the help of sound, as suggested in [19] and [20].

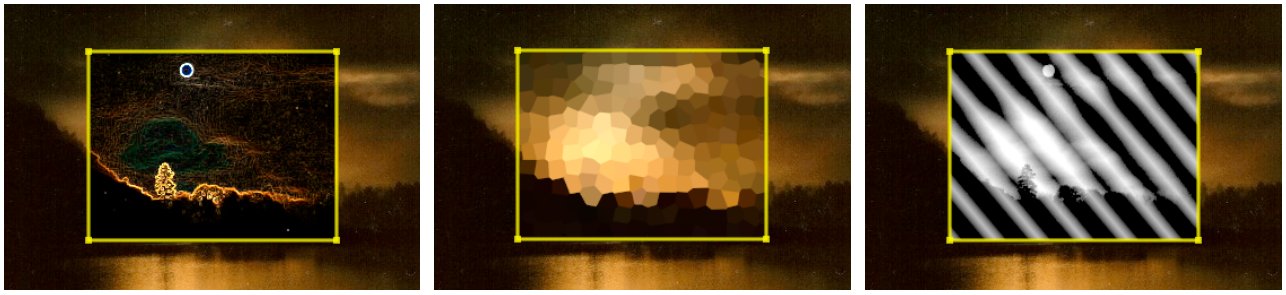
Data probing can be achieved by Rasterpiece, too. The size and the aspect ratio of a marquee are set to contain enough pixels to characterize the locality of selection, while being small enough to avoid any redundancy at the same time. This marquee can then be moved freely over the whole region to sonify the currently selected area. In fact, this is a combination of global probing and local scanning. Figure 9 shows an example of probing medical imaging data with Rasterpiece.

Although the size of probing marquee may vary, it is typically equivalent to less than 0.1[s] in duration and often shorter than the threshold of pitch/timbre perception, as mentioned in section 1.1. Consequently, sonified results of small probing marquees are mostly looped in order to provide meaningful auditory displays.

Looped sonification of small images can also occur under different circumstances, as described below in section 4.3.



(a) Original image. (b) Blur. (c) Posterize.



(d) Edges. (e) Crystallize. (f) Line screen.

Figure 8. Examples of filtered marquee. (a) is the original image, while the rest are filtered results (with different image filters, as specified respectively.)

4.2. Sound synthesis from images

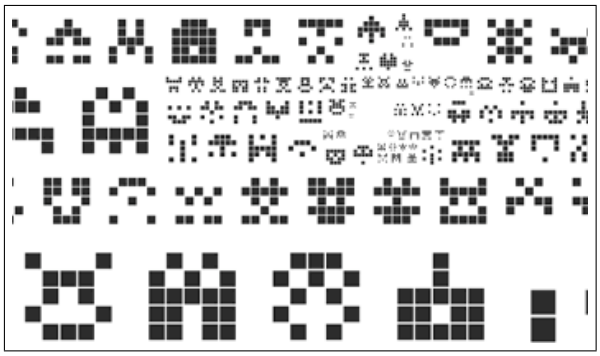
Perhaps the prime significance of Rasterpiece is that it allows for the use of almost every image as sound sample data, and that the image and its resulting sound are related in an evocative and predictable manner based on the raster sonification mapping. By selecting a specific region of an image and applying certain visual filters, the user can obtain the desired timbre and sonic texture as well as the basic auditory features (i.e., nominal pitch, additional pitch variation, and duration) of a sound. This leads to the idea of constructing a sound library from a collection of images. Figure 10 shows two sets of image patterns that have been used as sound sample libraries for Rasterpiece.

As a special case, *rastrogram* - raster-visualized image of a sound - can be sonified, too. Thanks to the reversible, one-to-one nature of raster mappings, this sonified result becomes a perfect "reconstruction" of the original sound by default, or can be a new sound if the rastrogram is manipulated in the visual domain.

4.3. Pulsar synthesis and sound particle design

Pulsar synthesis is a sound synthesis method based on the generation of sonic particles called *pulsars* [12] [13]. It can produce either rhythms or tones as it traverse perceptual time spans.

Figure 11 illustrates a pulsar, which contains an arbitrary waveform (called *pulsalet*) with duration d followed by a silent interval s. Repetitions of this signal produce a pulsar "train", whose timbre can be controlled by using different pulsalet waveforms and/or varying the lengths of



(a) Square patterns.

(b) Complex "sand-stroke" pattern.

Figure 10. Examples of visual patterns that are used as sound sample data by Rasterpiece. Images are obtained from [16].

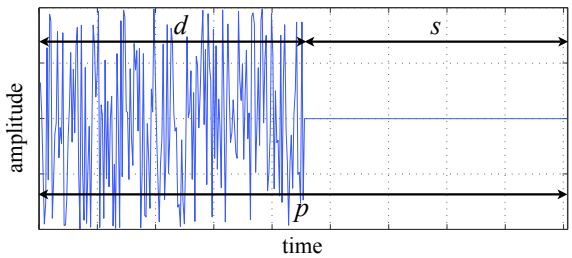
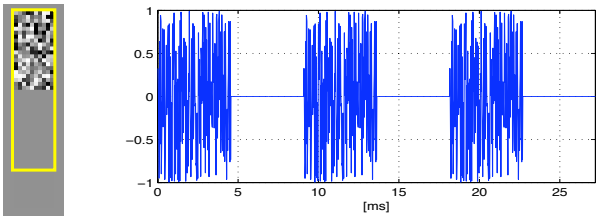
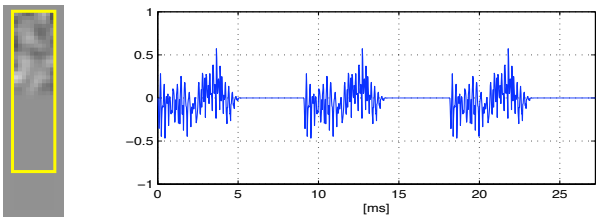


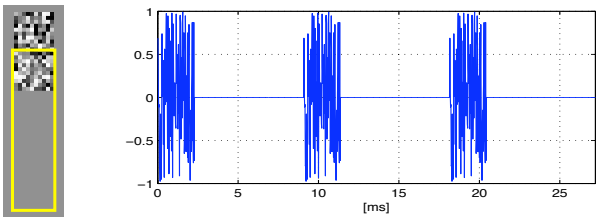
Figure 11. Example of a pulsar, with a white noise pulsalet of duration d followed by a silent interval of s. Overall length of the pulsar is p.



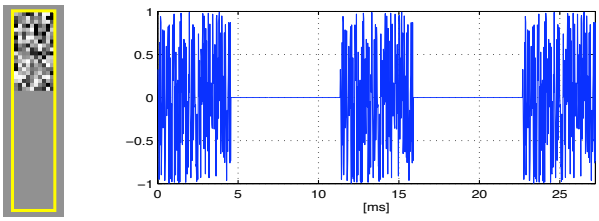
(a) White noise pulsalet. Size of the marquee is 10 x 40 pixels.



(b) Blurred white noise pulsalet, with the same marquee as in (a).



(c) White noise pulsalet, with shorter pulsalet duration. The marquee has the same size, but is on a different location.



(d) White noise pulsalet, with longer silent interval. Size of the marquee is 10 x 50 pixels.

Figure 12. Examples of different marquee selections (left) and their corresponding pulsar trains (right). Pitch of each result is determined by the size of the marquee (hence the overall duration of one pulsar): (a), (b), and (c) corresponds to 110.25 [Hz], while (d) is equivalent to 88.2 [Hz]. Sounds are sampled at 44,100 [kHz].

d and s. This effect is called as *pulsalet-width modulation*.

Rasterpiece enables the user to perform pulsar synthesis by manipulating the size and location of a small marquee and looping the sonified result. Figure 12 shows different marquees and their corresponding pulsar trains: in figure 12(b), the content of the original pulsalet (figure 12(a)) is blurred, whereas figure 12(c) is an example of a shorter pulsalet with the same overall duration p. Figure 12(d) illustrates the case of longer s with the same pulsalet as in figure 12(a). Obviously, variations in marquee settings produce different tonal characters.

If a pulsar is shorter than the threshold of perception, the pitch of its synthesized result is determined by p. Longer pulsars, on the other hand, produce infrasonic rhythms rather than tones: in this case, width of the current marquee (hence its nominal frequency) contributes to the perceived pitch, as mentioned before.

PulsarGenerator, an interactive sound synthesis program by de Campo and Roads, allows the user to design the waveshape and length of pulsalets and manipulate the pulsar's period to produce corresponding changes in timbre [7]. Compared with PulsarGenerator, Rasterpiece lacks a number of advanced features including control of modulation parameters and waveform display, but provides an image interface as a palette for intuitive pulsalet design.

5. CONCLUSION

Rasterpiece is a new multimedia software that combines image and sound based on the raster sonification rule. The cross-modality of raster mapping lays a solid foundation for intuitive integration of audio and visual information. Furthermore, interactive image processing and sound synthesis capabilities of Rasterpiece has strong implication for applications in various areas ranging from music and multimedia art to medical imaging.

Future works will include supporting more control options. In addition to computer keyboard and mouse, external devices (e.g., keyboards, controllers, physical interfaces, etc.) will be available for use with Rasterpiece via MIDI and Open Sound Control (OSC). This will enhance the playability and expressivity of Rasterpiece as a musical instrument or a performance tool.

A Real-time sound visualization module based on the raster mapping will be implemented, too. Together with the current sonification feature, this will complete the cycle of cross-modal conversions, thereby enabling a long chain of image sonification and sound visualization using digital filters in both audio and visual domains.

We also plan to construct a library of image-sound pairs that is perceptually meaningful and musically useful as well. This goes beyond the simple idea of using images as sound samples (or vice versa): we aim to develop it into a research on *synthetic synesthesia* with focus on musical applications, for which Rasterpiece will become the core experimental tool.

6. ONLINE RESOURCE

Rasterpiece is available for download at:
http://ccrma.stanford.edu/∼woony/software/rasterpiece/.

7. REFERENCES

[1] Apple Inc. "Core Image: ultra-fast, pixel-accurate", *http://www.apple.com /macosx/features/coreimage/*.

[2] Arfib, D., Dudon, J., and Sanchez, P. "WaveLoom, logiciel d'aide la cration de disques photosoniques", *http://recherche.ircam.fr/equipes/repmus/jim96 /actes/arfib/waveloom.html*.

[3] Berger, J. and Tal, O. B. "De natura sonoris", *Leonardo Music Journal*, 37(3), 2004.

[4] Bischoff, J., Gold, R., and Horton, J. "Music for an interactive network of microcomputers", *Computer Music Journal*, 2(3), pp.2429, 1978.

[5] Borgonovo, A. and G. Haus. " Musical sound synthesis by means of two-variable functions: Experimental criteria and results", *Proceedings of the International Computer Music Conference*, pp.3542, 1984.

[6] Cook, P. and Scavone, G. "The Synthesis Toolkit in C++", *http://ccrma.stanford.edu/software/stk/*.

[7] CREATE. "PulsarGenerator, a New Sound Synthesis Program for MacOS", *http://www.create.ucsb.edu/PulsarGenerator/*

[8] IRCAM. "AudioSculpt", *http://forumnet.ircam.fr/*.

[9] Jones, W. "Sight for Sore Eyes", *IEEE Spectrum*, February 2004.

[10] McLaren, N. and Jordan, W. "Notes on animated sound", *The Quarterly of Film, Radio, and Television*, 7(3), pp.223229, Spring 1953.

[11] Roads, C. "The Computer Music Tutorial", MIT Press, Cambridge, MA, USA, 1997.

[12] Roads, C. "Sound Composition with Pulsars", *Journal of The Audio Engineering Society*, 49(3), pp.134-147, 2001.

[13] Roads, C. "Microsound", MIT Press, Cambridge, MA, USA, 2002.

[14] Sedes, A., Courribet, B., and Thiebaut, J. "From the visualization of sound to real-time sonification: different prototypes in the Max/MSP/Jitter environment", *Proceedings of the International Computer Music Conference*, Miami, Florida, USA, 2004.

[15] U&I Software. "Metasynth 4", *http://uisoftware.com/MetaSynth/*.

[16] Tarbell, J. "Gallery of Computation", *http://complexification.net/*.

[17] Verplank, B., Mathews, M., and Shaw, R. "Scanned synthesis", *Proceedings of the International Computer Music Conference*, September 2000.

[18] Xenakis, I. "Formalized Music", Revised edition, Pendragon Press, New York, 1992.

[19] Yeo, W., Berger, J., and Wilson, S. "A Flexible Framework for Real-time Sonification with SonART", *Proceedings of the International Conference on Auditory Display*, Sydney, Australia, 2004.

[20] Yeo, W., Berger, J., and Lee, Z. "SonART: A Framework for Data Sonification, Visualization, and Networked Multimedia Applications", *Proceedings of the International Computer Music Conference*, Miami, Florida, USA, 2004.

[21] Yeo, W. and Berger, J. "Application of image sonification methods to music", *Proceedings of The Internaional Computer Music Conference*, Barcelona, Spain, 2005.

[22] Yeo, W. and Berger, J. "Raster Scanning: A New Approach to Image Sonification, Sound Visualization, Sound Analysis And Synthesis", *Proceedings of The International Computer Music Conference*, New Orleans, USA, 2006.

[23] Yeo, W. and Berger, J. "Application of Raster Scanning Method to Image Sonification, Sound Visualization, Sound Analysis and Synthesis", *Proceedings of The International Conference on Digital Audio Effects*, Montreal, Canada, 2006.

[24] Yeo, W. "StkX - an STK framework for OS X", *http://ccrma.stanford.edu /∼woony/software/stkx/*