sound/tracks:
Real-Time Synaesthetic Sonification of Train Journeys

Peter Knees, Tim Pohle, and Gerhard Widmer
Department of Computational Perception
Johannes Kepler University Linz, Austria
peter.knees@jku.at, tim.pohle@jku.at, gerhard.widmer@jku.at

ABSTRACT

Travelling on a train and looking out of the window at the moving scenery reveals a composition of "visual music" with its own tempo and rhythm, its own colours and harmonies. The project sound/tracks aims at capturing these visual impressions and translates them into a musical composition in real-time - producing an immediate and unique soundtrack to the train journey based on the passing landscape. To this end, the outside impressions are captured with a camera and translated into instantaneously played back piano music. The immediately added sound dimension allows for reflection of the visual impression and deepening of the state of contemplation. For the resulting compositions, the passing scenery can be considered the score. "Re-transcription" of this score to an image gives a panoramic overview over the complete journey and exhibits some interesting effects caused by the movement of the train, such as compression and stretching of passing objects. In addition to intensifying the experience of a train journey, sound/tracks permits to persistently capture and archive the fleeting impressions of journey and composition and allows for re-experiencing the trip both visually and acoustically at a later point.

Figure 1: sound/tracks on a PC/laptop. The video is displayed in real time in the top panel, the condensed history of the journey slowly moves past in the form of a 'score image' in the lower panel.

Categories and Subject Descriptors

J.5 [**Computer Applications**]: Arts and Humanities

General Terms

Algorithms

1. MOTIVATION AND RELATED WORK

When sitting on a train, many people enjoy listening to music and – often simultaneously – looking out of the window and watching the landscape passing by. However, in most cases, the impressions from these two tasks do not correspond with each other, i.e., visual stream and sound stream remain separate entities with no aesthetic connection. With the project presented here, we aim at developing an artistic application called sound/tracks that bridges and brings together these entities by creating a synaesthetic experience where visual and acoustic impressions correspond to each other. To this end, the images on the outside of the train are captured with a video camera and translated into (piano) music that complements the visual impression in real-time. This allows for a reflection of the fleeting visual impressions and deepening of the state of contemplation. The passing scenery can be considered the score of a musical composition which is going to be interpreted based on outside conditions such as weather and lighting, the speed of the train, the quality of the camera, and the degree of staining of the window glass. Thus, every journey will produce a unique composition. In addition to intensifying the experience of a train journey, sound/tracks should be usable on-line (i.e., while travelling) as well as off-line (e.g., for public exhibits or private re-experiencing of trips).

A major inspiration for this work was the music video for the track "Star Guitar" by "The Chemical Brothers" directed by Michel Gondry [4]. The video gives the impression of a continuous shot filmed from a passenger's perspective on a speeding train. The train passes through visually rich towns, industrial areas, and countryside where all buildings and objects passing by appear exactly in sync with the various beats and musical elements of the track. While in that video the visual elements were composed based on the musical structure, we aim at achieving the opposite in our project: our goal is to give the passing landscape an active role to control the real-time production of music. In our devised sonification algorithm, the translation of light and colour to sounds refers back to the Russian composer,

MM'08, October 26–31, 2008, Vancouver, British Columbia, Canada.
ACM 978-1-60558-303-7/08/10.

Figure 2: A visual 'score image'. Acceleration of the train can be observed from left (no movement, i.e. stop at a station) to right: with increasing speed, initially "stretched" objects become more and more compressed.

pianist, and (self-claimed) synaesthete Aleksandr Skrjabin. Skrjabin created a mapping between piano tonalities and colours based on his visual (colour) sensations when hearing music in a specific key [8]. Beside this historically motivated guidance, we also aimed at incorporating findings from psychology into our algorithm. From the literature, we could find indications that there are associations of pitch height with all of brightness [2], colour hue [3], and spatial height [7]. For an application like the one we are building, the crucial question is how these can be combined into an overall function that maps visual impressions into sounds.

2. TECHNICAL REALISATION

In the current state of sound/tracks, translation of visual input to acoustic output is performed as follows: The data captured by the camera is given as a series of images (frames), each represented as an array of *rgb* pixels. The captured frames are analysed at a constant rate of 7 frames per second, providing the dominant rhythm of the composition. From each analysed frame, we take the middle column of the pixels and divide it into 4 parts of equal height. Each of these parts then is used to generate tones played in a different octave. The pixels of a part are transformed to a pitch by calculating the cosine distance of each contained pixel's value in the *hsv* colour space to all of the twelve colours on Skrjabin's colour keyboard (cf. [8]). These values then are subsumed into a twelve-binned histogram over all pixels. The fullest (and if above a threshold, also the second fullest) bin is the pitch that is played in this octave. To avoid repetitively playing the same notes at similar subsequent frames, similar notes are held for at most 7 frames (1 sec.). In some cases, this results in repetitive patterns that are perceived as musical themes. The corresponding notes are played on a MIDI sound generator with a piano sound. A more detailed description of the algorithm can be found in [5].

Due to the steady rate of 7 fps, there is a clearly noticeable basic rhythm pattern in the music, which the listener may associate with the steady progression of the train. The resulting harmonies are quite pleasurable, which might be a result of the colour distribution in the mapping from colours to pitches. Also, a changing landscape is reflected in the resulting music, while the overall feeling remains the same. The video available at the project web site [1] gives a more direct impression.

3. PANORAMIC SCORE IMAGES

"Re-transcription" of the analysed columns to another image (by sequentially adding the processed columns along the x-axis) yields a panoramic overview over the scenery passed so far. Such a panoramic image captures and displays the *dynamics* of the journey and exhibits some interesting effects caused by the movement of the train. Since frame rate and position of the camera are both static, proximity of objects and slope and velocity of the train result in characteristic visual effects. For example, objects filmed at high speeds are displayed very narrow, whereas objects filmed at low speeds appear stretched (cf. Figure 2).[1]

4. PRESENTATION

sound/tracks can be used on-line, during a journey, both on laptop and on appropriate mobile phones. We plan to make the software freely available to the general public via our project web page [1]. Train journeys that have been recorded at high quality can be adapted for performances in public space, e.g. as a multimedia exhibit or even as a visually augmented piano concert. In such an environment, sonification could be performed live and in high-quality on a computer-controlled grand piano. To complement the audio/video aspect of the exhibit, entire journeys in the form of printouts of score images could be glued around the walls of the exhibition room to document train journeys off-line.

5. ACKNOWLEDGMENTS

Research underlying this work is supported by the Austrian Fonds zur Förderung der Wissenschaftlichen Forschung under project numbers L112-N04 and L511-N15.

6. REFERENCES

[1] sound/tracks Project Web Site. URL: `http://www.cp.jku.at/projects/soundtracks/`

[2] W. G. Collier and T. L. Hubbard. Musical scales and brightness evaluations: Effects of pitch, direction, and scale mode. *Musicae Scientiae*, 8:151–173, 2004.

[3] D. L. Datteri and J. N. Howard. The sound of color. In *Proc 8^{th} Int'l Conference on Music Perception & Cognition*, Evanston, USA, 2004.

[4] M. Gondry. Music video for "The Chemical Brothers – Star Guitar", 2002.

[5] T. Pohle and P. Knees. Real-Time Synaesthetic Sonification of Traveling Landscapes. In *Proc 5^{th} Int'l Mobile Music Workshop*, Vienna, Austria, 2008.

[6] M. Reinhart and V. Widrich. tx-transform. URL: `http://www.tx-transform.com/` (last access: 2008/04/28).

[7] E. Rusconi, B. Kwan, B. L. Giordano, C. Umiltà, and B. Butterworth. Spatial representation of pitch height: the SMARC effect. *Cognition*, 99:2:113–129, 2006.

[8] Wikipedia. Alexander Scriabin. URL: `http://en.wikipedia.org/wiki/Alexander_Scriabin` (last access: 2008/04/28).

[1] Note that similar effects (e.g. stretching and compression of objects) can also be observed using the *tx-transform* technique by the artists Martin Reinhart and Virgil Widrich [6].