

Organ Augmented Reality: Audio-Graphical Augmentation of a Classical Instrument

Christian Jacquemin, LIMSI-CNRS, France

Rami Ajaj, LIMSI-CNRS, France

Sylvain Le Beux, LIMSI-CNRS, France

Christophe d'Alessandro, LIMSI-CNRS, France

Markus Noisternig, IRCAM, France

Brian F. G. Katz, LIMSI-CNRS, France

Bertrand Planes, Artist, France

ABSTRACT

This paper discusses the Organ Augmented Reality (ORA) project, which considers an audio and visual augmentation of an historical church organ to enhance the understanding and perception of the instrument through intuitive and familiar mappings and outputs. ORA has been presented to public audiences at two immersive concerts. The visual part of the installation was based on a spectral analysis of the music. The visuals were projections of LED-bar VU-meters on the organ pipes. The audio part was an immersive periphonic sound field, created from the live capture of the organ's sound, so that the listeners had the impression of being inside the augmented instrument. The graphical architecture of the installation is based on acoustic analysis, mapping from sound levels to synchronous graphics through visual calibration, real-time multi-layer graphical composition and animation. The ORA project is a new approach to musical instrument augmentation that combines enhanced instrument legibility and enhanced artistic content.

Keywords: Augmented Musical Instrument, Augmented Reality, Organ Augmented Reality (ORA), Real-Time Visualization, Sound to Graphics Mapping

INTRODUCTION

Augmented musical instruments are traditional instruments that are modified by adding controls and additional outputs such as animated

graphics (Bouillot et al., 2009; Thompson et al., 2007). The problem with usual approaches to instrument augmentation is that it generally makes the instrument more complex to play and more complex to understand by the spectators. The enhanced functionality of the instrument often distorts the perceived link between the

DOI: 10.4018/jcicg.2010070105

performer's actions and the resulting sound and images. Augmentation is likely to confuse the audience because it lacks transparency and legibility.

In addition to augmenting traditional instruments with new controllers, like the hyper-kalimba (Rocha et al., 2009) which extends the kalimba (an instrument from the percussion family), augmented reality is also used to create new musical instruments. Some of these instruments mimic real music devices like the Digital Baton (Marrin et al., 1997), replicating the traditional conducting baton, or the AR scratching¹ imitating a DJ's vinyl scratch. Other musical instruments that use augmented reality are totally innovative and are not based on existing devices. The Augmented Groove (Poupyrev et al., 2001) is an example of such a device where novice users manipulate a physical object in space to play electronic musical compositions. The main difference between creating novel instruments and extending existing instruments is the level of familiarity with the instrument. Instrument extension seems more suitable for experimented performers rather than novice ones due to the experience level with the instrument and possibly a wider range of control.

Musical instrument augmentation is interesting because it extends a traditional instrument, while preserving and enriching its performance and composition practices. The Organ and Augmented Reality (ORA) project focuses on a rarely stressed use of augmentation, the enhanced comprehension and legibility of a music instrument without increasing its complexity and opacity. Our research on output augmentation follows the same purposes as (Jordà, 2003), making the complexity of music more accessible to a larger public. Jordà's work focused on the playing experience; similarly, we intend to improve and facilitate the listening experience. These principles have been used by Jordà et al. (2007) for the design of the ReacTable, an augmented input controller for electronic musical instruments. The ReacTable is a legible, graspable, and tangible control interface, which facilitates the use of an electronic instrument so as to be

accessible to novices. Its use by professionals in live performances confirms that transparency is not boring and is compatible with long term use of the instrument.

This paper presents the issues and technical details of the ORA project and performance, the augmentation of an historical church organ for a better understanding and perception of the instrument through intuitive visual and audio outputs. It is based on the following achievements:

- The visuals are directly projected onto the organ pipes (not on peripheral screens),
- The visual augmentation is temporally and spatially aligned: the visual rendering is cross-modally synchronized with the acoustic signal and the graphical projection is accurately aligned with the organ geometry,
- The augmentation preserves the traditional organ play. Traditional compositions as well as new artworks can be played on the augmented instrument,
- The augmentation offers a better understanding of the instrument's principles by showing a visualization of hidden data such as the spectral content of the sound and its position inside the instrument.

The aim of the ORA project was to make an audio and visual augmented reality on the grand organ of the Sainte Elisabeth church in Paris. The ORA project was supported by the City of Paris "Science sur Seine" program for bringing science closer to citizens. The pedagogical purpose was to present the basic principles of sound and acoustics, and illustrate them through audio and graphics live performances. The two concerts were complemented by a series of scientific posters explaining background knowledge and specialized techniques used in the ORA project. The project involved researchers in interactive 3D graphics and computer music, a digital visual artist, an organ player and composer, and engineers.² ORA has been presented to public audiences through two visually and acoustically augmented concerts at Church Ste Elisabeth.

Figure 1. ORA Concerts, Eglise Sainte Elisabeth de Hongrie, Paris, May 15th & 17th, 2008



The visual part of the installation was based on a spectral analysis of the music. The visuals were projections of LED-bar VU-meters on the organ pipes. The audio part was an immersive periphonic sound field, created from the live capture of the organ sound, so that the listeners had the impression to be placed inside the augmented instrument. This article presents in detail the visual augmentation; the audio part of this project is described in (d'Alessandro et al., 2009).

VISUAL AUGMENTATION OF INSTRUMENTS

Musical instrument augmentation can target the interface (the performer's gesture capture), the output (the music, the sound, or non-audio rendering), or the intermediate layer that associates the incoming stimuli with the output signals (the mapping layer). Since the ORA project approach tries to avoid modifying the instrument's playing techniques, it focuses on the augmentation of the mapping and output layers in order to enhance composition, performance, and experience.

About augmented music composition, Sonofusion (Thompson et al., 2007) is both a programming environment and a physically-augmented violin used for composing and performing multimedia artworks. Sonofusion compositions are "written" through lines of code, and the corresponding performances are controlled in real-time through additional knobs, sliders, and joysticks on the violin. While addressing the question of multi- and cross-modal composition and performance in a relevant way, Sonofusion control system is complex and opaque. The many control devices offer multiple mapping combinations. Because of this diversity, the correlation between the performer's gestures and the multimedia outputs seems arbitrary to the audience at times. Musikalscope (Fels et al., 1998) is a cross-modal digital instrument that was designed with a similar purpose, and has been criticized by some users for its lack of transparency between the user's input and its visual output.

For teaching music to beginners, augmented reality can be used to project information onto the instrument about the playing of the instrument. For electric and bass guitar,

Cakmakci et al. (2003) and Motokawa and Saito (2006) augment the instrument by displaying the expected location of the fingers onto the guitar fingerboard. The visual information is synchronized with audio synthesis and is available through direct projection on the instrument or via visual composition using a head-mounted display.

About augmenting a performance, the Synesthetic Music Experience Communicator (Lewis Charles Hill, 2006) (SMEC) is used to compose synesthetic cross-modal performances based on visual illusions experienced by synesthetes. When compared with Sonofusion, graphic rendering in SMEC is better motivated because it relies on reports of synesthetic illusions. SMEC however raises the question whether we can display and share perceptions which are deeply personal and intimate in nature.

Visually augmented performances have also addressed human voice augmentation. The *Messa di Voce* installation (Levin & Lieberman, 2004) was designed for real-time analysis of human voice and for producing visual representations of human voice based on audio processing algorithms. The graphical augmentation of the voice was autonomous enough to create the illusion of being an alter ego of the performer. Since it is governed by an “intelligent” program, *Messa di Voce* graphical augmentation does not seem as arbitrary as other works on instrument augmentation (human voice is considered here as an instrument). When attending augmented instrument performance with insufficiently motivated augmentation, the spectators are immersed by a complex story which they would not normally expect when attending a musical event.

VIRTUAL AND AUGMENTED REALITY FOR THE ARTS AT LIMSI-CNRS

In 2003, LIMSI-CNRS launched a research program entitled *Virtuality, Interactivity, Design, and Art (VIDA)* to develop joint projects with artists, designers, and architects in Virtual and Augmented Reality, and, more generally,

in arts/science projects on Human/Computer Interaction. Through artistic collaborations, new research themes have emerged which have fertilized research at LIMSI-CNRS and broadened the scope of our works. The developments enabled by these academic works have also provided artists with new software tools and new environments for their creative works. Without these developments, they would not have been able to realize such innovative artworks. A necessary engineering workforce has been involved in these collaborations so that scientific prototypes could be turned into usable applications for the artists whether on stage or in art installations.

The ORA project is part of a sub-theme in VIDA dealing with Augmented Reality in the arts. This theme was initiated through collaboration with the theater company Didascalie. Net (director Georges Gagneré) concerning video-scenography for the performing arts. The question of “presence” in an artwork, smart projection on non-flat surfaces, and the interaction of performers with live image synthesis were among the issues addressed in this collaboration. Through live “experiments” with the stage director and actors, unexpected experimental configurations emerged which triggered innovative works. For instance, the combination of video-projection and conventional lighting raised new topics of research considering video-projection and performer’s or spectator’s shadows. Interaction of performers with live computer graphics has also led us to develop a dynamic multilayer model for video-scenography that parallels the layers of stage decoration (Jacquemin & Gagneré, 2007).

Augmented Virtuality (closer to the digital world than Augmented Reality) has also been used in collaborations between the visual artist Bertrand Planes and LIMSI-CNRS for two digital art installations (*Mar:3D* and *Gate:2.5*) in which shadows of spectators were projected into the virtual scene (Jacquemin et al., 2007). Through these installations, we have addressed the issue of spectator presence in a Virtual Environment through the use of shadows, which also proved to be a good medium for non-tactile gestural exploration of a virtual world.

The ORA project was developed in 2008 to address issues of accurate spatial and temporal registration of video-projection in a real-time performance. It was the first time that LIMSI-CNRS was involved in a project with strong expectations for aligning real time graphics in space with a complex architecture and aligning them in time with a live sound production. This project has been followed since by works on Mobile Augmented Reality³, in which the viewers' location in the scene changes with time. The artistic target of this work was an installation on the River Seine, for which spectators embarked on a boat cruise, viewing the river banks augmented with a re-projection of modified infrared video-capture of the riverside. As a first approach to mobile augmented reality, we have not dealt with the identification of mobile elements but only with the issues of dynamic calibration and live special effects. Future work will deal with more elaborate analysis of the mobile scene related to tracking and identification in the physical world allowing for semantic registration of virtual elements on the real-world.

ORA ARTISTIC DESIGN

Visual Design. The design of the ORA project visual artwork transforms the instrument in such a way that it appears as both classical and contemporary. The 20th century style of the visual augmentation contrasts with the baroque architecture of the instrument. The Sainte Elisabeth church organ is located high on the rear wall of the building. In this church, as it often occurs, the believers face the altar and listen to the music and cannot look at the instrument. Even if one looks at the organ, the organist cannot be seen, resulting in a very static visual experience. For the two ORA project concerts the seating was reversed, with the audience facing the organ at the gallery.

The church acoustics is an integral part of the organ sound as perceived by the listeners. Through the use of close multichannel microphone capture, rapid signal processing, and a multichannel reproduction system, the audience was virtually placed inside a modified "organ" acoustics, thereby providing a unique sound and music experience. To highlight the digital augmentation of the organ sound, VU-meters were displayed on the pipes of the organ facade through videoprojection, making a common reference to audio amplifiers. These VU-meters were dynamically following the music spectral composition and built a subtle visual landscape that was reported as "hypnotic" by some members of the audience. The traditionally static and monumental instrument was visually transformed into a fluid, mobile, and transparent contemporary art installation.

Sound Effects & Spatial Audio Rendering.

The organ is one of the oldest musical instruments in Western musical tradition. It offers a large pitch range with high dynamics, and it can produce a large variety of timbres. An organ is even able to imitate orchestral voices. Because of the complexity of the pipe machinery, advances in organ design have been closely related to the evolution of associated technologies. During the second half of the 20th century electronics were applied to organs for two purposes:

- To control the key and stop electro-pneumatic mechanisms of the pipes,
- To set the registration (the combinations of pipe ranks).

However, very little has been achieved for modifying the actual sound of the organ. In the ORA project, the pipe organ sound is captured and processed in real-time through a sequence of digital audio effects. The transformed sound is then rendered via an array of loudspeakers surrounding the audience. Therefore, the sound perceived by the audience is a combination of the natural sound of organ pipes, the processed sound, and the room acoustics related to each of

these acoustic sources. Spatial audio processing and rendering places the inner sounds of the organ in the outer space, radically changing their interaction with the natural room acoustic, and adding a new musical dimension.

Augmented Organ. Miranda and Wanderley (2006) refer to augmented instruments⁴ as “the original instrument maintaining all its default features in the sense that it continues to make the same sounds it would normally make, but with the addition of extra features that may tremendously increase its functionality”. With a similar intention in mind, the ORA project was designed to enrich the natural sound of organ pipes through real-time audio processing and multi-channel sound effects, to meet the requirements of experimental music and contemporary art.

ARCHITECTURE AND IMPLEMENTATION

The Instrument

The Sainte Elisabeth Church organ used for the ORA project is a large 19th century instrument (a protected historical monument), with three manual keyboards (54 keys), a pedal board (30 keys), and 41 stops with mechanical action. This organ has approximately 2500 pipes. Only 141 of the organ pipes are located on the facade and visible to the public. The front side of the organ case has a dimension of approximately 10x10m. The organ pipes are organized in four main divisions: the “Positif”, a small case on the floor of the organ loft (associated with the first manual keyboard), the “Grand Orgue” and “Pédale” divisions at the main level (associated with the second manual keyboard and the pedal board), and the “Récit” division, a case of about the same size as the “Positif”, crowning the instrument (associated with the third manual keyboard). The “Récit” is enclosed into a swell-box.

A set of 5 microphones was placed in the four instrument divisions (see Figure 2 left). These divisions are relatively sound isolated, and the near-field sound captured in one division was significantly louder than the sounds received from other ones. Hence, the captured sounds can be considered as being acoustically “isolated” from each other at least in the mid- and high-frequency ranges.

General Architecture

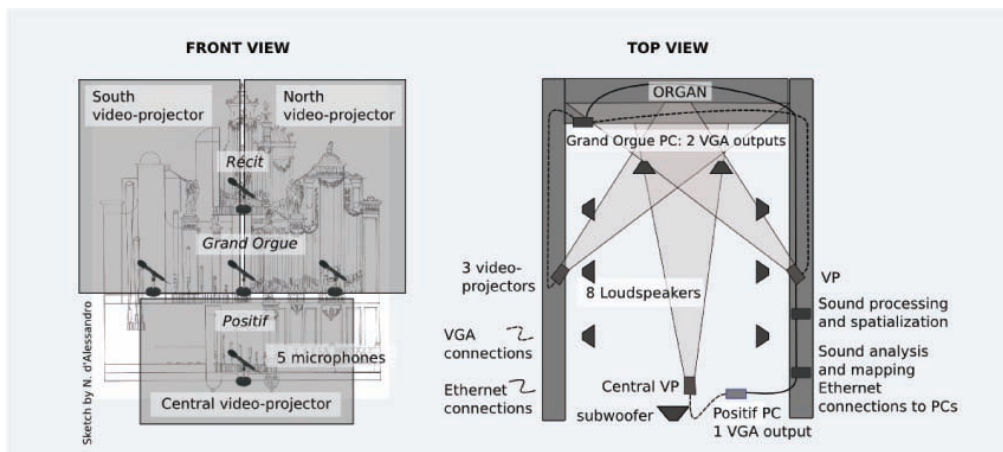
The organ pipes, despite their gray color and their slight specular reflection, were an appropriate surface for video-projection. The visual ornamentation of the instrument was made with three video-projectors: two for the upper part of the instrument and one for the lower part (see Figure 2).

The organ sound captured by the microphones was given as input to a digital signal processing unit for sound analysis and special effects. The processed sounds were then diffused back into the church over a loudspeaker array encircling the audience. The sound processing modules for graphical effects consisted of spectral analysis and sampling modules that computed the levels of the virtual VU-meters. These values were sent over the Ethernet network to the 3D engine. The graphical rendering relied on Graphic Processing Unit programming: vertex and fragment shaders that used these values as parameters to animate the textures projected on the organ pipes to render the virtual LED-bars. The right part of Figure 2 shows the full hardware installation of the ORA project with the location of the video-projectors and loudspeakers, and the main data connections.

Graphic Rendering

Graphic rendering relies on Virtual Choreographer (VirChor)⁵, a 3D graphic engine offering communication facilities with audio applications. The implementation of graphic rendering in VirChor involved the development of a calibration procedure and dedicated shaders for blending, masking, and animation. The architecture is divided into three layers:

Figure 2. Architecture of the installation: sound capture and video-projection



initial calibration, real-time compositing, and animation.

Calibration. The VU-meters are rendered graphically as quads that are covered by two samples of the same texture (white and colored LED-bars), depending on the desired rendering style. These quads must be registered spatially with the organ pipes. Due to the complexity of the instrument and its immobility, registration of the quads with the pipes was performed manually. Before the concert began, a digital photograph of the projection of a white image was taken with a camera placed on each video projector, near the projection lens. This photo was then used as a background image in Inkscape⁶ to calibrate the projection of the virtual LED-bars on the organ pipes. The vector image in Inkscape contained as many quads as visible organ pipes in the background image. Each quad was manually aligned with the corresponding pipe of the background image in the vector image editor. Luckily, the amount of effort for this calibration work was only significant for the first edition of the vector image. Successive registrations (for each re-installation) amounted to a

slight translation of the quads aligned in the previous edition, since attempts were made to relocate each video-projector in a similar position for each concert. The resulting Inkscape SVG vector image was then converted into an XML 3D scene graph through a Perl script and then loaded into VirChor.

During a concert, the VU-meter levels were received from the audio analysis component (section *Analysis and Mapping*) and transmitted to the Graphic Processing Unit (GPU) that in turn handled the VU-meter rendering. GPU programming offered a flexible and concise framework for layer compositing and masking through multi-texture fragment shaders, and for interactive animation of the VU-meters through vertex shader parameterization. Moreover, the use of one quad per VU-meter per visual pipe handled by shaders facilitated the calibration process. Frame rate for graphic rendering was above 70 FPS and no lag could be noticed between the perceived sound and the rendered graphics.

Compositing. The graphical composition was organized into 4 layers: (1) a background layer made of a quad that contained an image of the organ pipes, (2) an animated

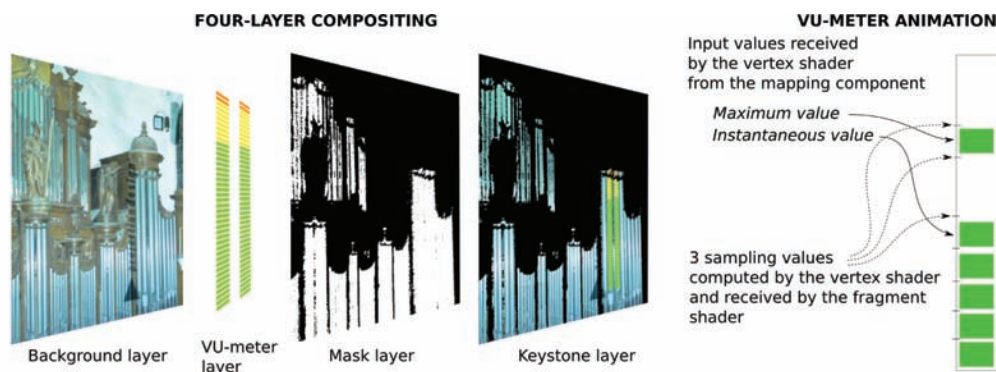
layer made of as many quads as organ pipes, each quad was used to render one of the VU-meters, (3) a masking layer made of a single black and white mask, and used to avoid animated quads to be rendered outside the organ pipes, and (4) a keystone layer used to distort the output image and register it accurately on the organ pipes (see left part of Figure 3). The VU-meter animated layer is made of a set of multi-textured quads, and the background and mask layers are single quads that are parallel to the projection plane and that fill the entire display. Real-time compositing, keystone homography, and control of background color were made through vertex and fragment shaders applied on the geometrical primitives building these layers. The keystone layer (4) is a quad textured by the image generated by layers (1) to (3) that is oriented in such a way that it can correct the registration of the virtual VU-meters on the organ pipes. A modification of the keystone quad orientation is equivalent to applying an homography to the final image. This transformation enables slight adjustments in order to align the digital graphics with the organ and to compensate for calibration inaccuracies. It could also be computed automatically from real-time captures of calibration patterns (Raskar & Beardsley, 2001). Elaborate testing has shown that the background, VU-meter, and mask quads were perfectly registered with the physical organ, and thus made the keystone layer unnecessary. The masking layer is a quad textured with a black and white image of the organ facade where the pipes are in white and the organ wood parts are black. It is used to avoid any projection of the VU-meters onto the wooden parts of the organ, and also to apply a gray color on the part of the organ pipes that is not covered by VU-meters to make them possibly visible to the audience.

Animation. The animated VU-meter layer is made of textured quads registered with all the visible pipes of the organ. The texture for VU-meter display is made of horizontal colored stripes and a transparent background (42 stripes for each pipe of the Grand Orgue and Récit and 32 stripes for each pipe of the Positif). The purpose of the animation of these VU-meter quads is to mimic real LED-bar VU-meters that are controlled by the energy of their associated organ sound spectral band (see next section).

Each VU-meter receives activation values from the sound analysis and mapping components: the instantaneous value and the maximum value for the past 500ms (typical peak-hold function). These values are received through UDP messages and represent the sound levels of the spectral bands associated with each pipe. These intensities are sampled in the vertex shader to show or hide whole texture stripes and to avoid displaying only fractions of them. The level sampling performed in the vertex shader and applied to each quad is based on a list of predefined *sampling values* loaded in the shader. Since the height of a VU-meter texture is clamped to $[0, 1]$, each sampling value is the height between two stripes that represent two VU-meter bars (see right part of Figure 3). For example, a texture for 42 LED-bars has 43 sampling values. The sampling values are then transmitted from the vertex shader to the fragment shader that only displays the stripes below the received values and the top stripe associated with the maximal sampled value. The resulting perception by the audience is that each VU-meter is displaying a number of LED-bar stripes that corresponds to the associated spectral band intensity.

Before describing how these instantaneous control values were generated by sound analysis, we first present the detailed content of the musical program and its motivation.

Figure 3. Multi-layer composition and VU-meter animation through sampling



Musical Program

The ORA project was based on two organ concerts, with a bit of *strangeness* added by a virtual graphic animation of the organ facade and live electronic modifications of the organ sound. The musical program was a combination of a “classical” program and somewhat unusual digitally augmented pieces. The pieces of the great classical organ repertoire (Bach, Couperin, Franck, Messiaen) were alternated with a piece in 12 parts especially written by Christophe d’Alessandro for this project. This piece exploited the various musical possibilities offered by the sound capture, its digital transformation, and the diffusion system.

A large majority of special effects cannot be applied to the classical repertoire without damaging their subtle musical content. Contrary to the aesthetics of classical music played on electronic instruments, we adopted the point of view of historically informed performances, privileging historical registrations. Along this line, the music played in the concert was chosen to fit the aesthetics of the specific organ considered. Only subtle spatial audio and reverberation effects were used in conjunction with classical music.

It must be pointed out that the application of electronic effects to classical music is somewhat paradoxical: the effects in this case are considered successful as long as they do not

sound “electronic”, or other words, as long as they are not noticed by the audience.

The main argument of the musical piece composed for the ORA project in 12 parts was to play with inner and outer spaces, capturing inside and playing outside the instrument. This argument is also a metaphor for the music itself, based on a short text by Dorothée Quoniam: *Les 12 degrés du silence* (the 12 degrees of silence). Quoniam, a 19th-century Carmelite, explained to a young sister the teachings of her inner voice. The cycle is about speech, silence, inner and outer voices. It was played in alternation with classical repertoire music. This piece makes use of several unusual sound possibilities offered by the system.

1. *Sound relocation in the church.* The sound captured in a given division is played at another place in the church.
2. *Dynamic sound location.* Sound motion is suited to music made of an accompanied solo voice (called “Récit” in the French organ literature), like a singer moving in the church. A more massive effect is the slow extension and retraction through variation of the spatial extent (width) of the sound of a division in the acoustic space, like a tide rising and falling.
3. *Virtual room augmentation.* Artificial reverberation enlarges the acoustic space. This can transform the acoustics of the relatively small church where the concert

took place into that of a grand cathedral. At the same time, sounds presented to the audience through loudspeakers have less interaction with the natural room acoustics before arriving at the audience, resulting in a perceived reduction in reverberation.

4. *"Additive" effects that enrich the original sound.* Additive effects work well when applied to flute pipes, adding artificial harmonics to sound. For instance, the inharmonicity provided by the harmonizer effect and reverberations can transform the pipe sounds into percussion-like sounds.
5. *"Subtractive" effects that spectrally reshape the original sound.* Subtractive effects work well when applied to spectrally rich sounds. For instance the spectrum shaping effect provided by the Kapkus-Strong algorithm can give a vocal quality to reed pipes.

Analysis and Mapping

This section describes the real-time audio analysis and mapping for VU-meter visualization. Most of the approximately 2500 organ pipes are covered by the organ case, while only the 141 of the facade pipes are seen by the audience. As such, a direct mapping of the frequency played to visual pipes is not relevant, due to the large number of hidden pipes. In the context of ORA, the main purpose of the correspondence between audio data and graphical visualization was:

1. to metaphorically display the energy levels of the lowest spectral bands on the largest pipes (resp. display the highest bands on the smallest pipes)⁷,
2. to maintain the spatial distribution of the played pipes by separating the projected spectral bands in zones, corresponding to the microphone capture regions and thereby retaining the notion of played pipe location,
3. to visualize the energy of each spectral band in the shape of a classical audio VU-meter. The display was based on the instantaneous value of the energy and its last maximal value with a slower refreshing rate.

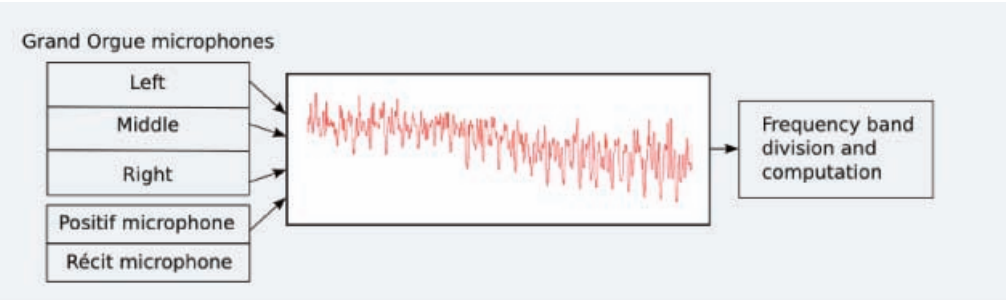
In order to estimate the mapping of sound level values to VU-meter heights, pre-recordings were analyzed (Figure 4). This analysis allowed for a rough estimate of the overall range of the various organ divisions and to separate these spectral ranges into different frequency bands according to the evolution of the harmonic amplitudes over frequency. The analysis resulted in a maximum spectral range of 16 kHz for the Positif and Récit divisions of the organ, and 12 kHz and 10 kHz for the central and lateral parts of the Grand Orgue.

Each spectral band was further divided into sub-bands corresponding to the number of visually augmented pipes, i.e. 33 for Positif and Récit, 20 and 35 for the lateral and central Grand Orgue. The sub-bands were not equally distributed over frequency range (warping) in order to gain a better energy balance between low and high frequencies. The spectral energy contained in the lowest frequency range was much greater than in the highest one. Thus, the frequency bands widths for lower frequencies were narrower, so as to have approximately the same spectral dynamics over all frequency bands. The frequency band divisions are summarized in Table 1. The energy of the lowest sub-band (the largest pipe) was used as reference signal for re-calibration.

The real-time spectral analysis consists of three stages: estimation of the power spectral density for each sub-band, mapping, and broadcasting over IP. The concert mapping process is presented schematically in Figure 5.

Power spectral density (PSD). The PSD was estimated via periodograms as proposed by Welch (1967). The buffered and windowed input signal was Fourier transformed (Fast Fourier Transform, FFT) and averaged over consecutive frames. Assuming ergodicity, the time average provides a good estimation of the PSD. Through long term averaging, the estimated sub-band levels are not sensitive to brief peaks and represent the root mean square (RMS) value. The decay of the recursive averaging was adjusted such

Figure 4. Spectral sound analysis



that the VU-meter values changed smoothly for visual representation. The actual averaging was such that every incoming frame was added to the last three buffered frames.

Frequency band division. The computed periodograms were transmitted to the frequency band division module as five 512-point spectra. This module, that represents the second part of sound signal processing, divided the Welch periodograms into 141 frequency bands. Since the number of visible pipes in each division of the organ was inferior to 512 (resp. 33, 20, 35, 20, and 33) an additional spectral averaging was necessary in order to map the entire frequency range to the pipes of the organ. According to the spectral tilt, lower frequency bands (below ~1.5 kHz) had more energy, thus only three frequency bands from the Welch periodograms were added for the largest pipes,

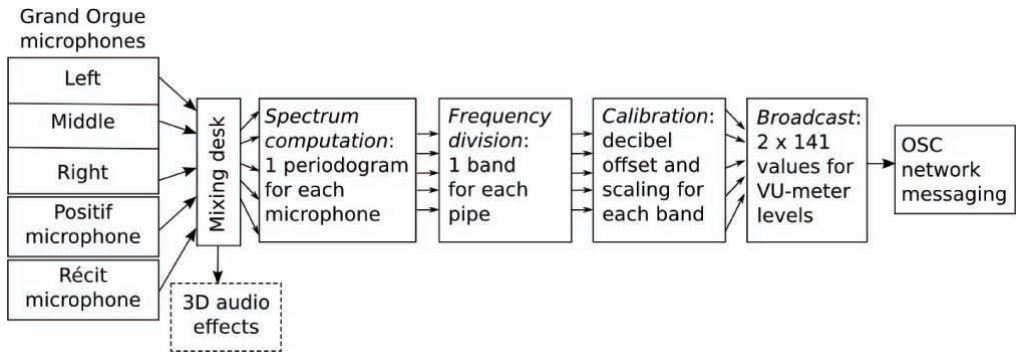
whereas up to 30-40 bands were added for the highest frequency range (above ~8 kHz), as detailed in Table 1. For the central Grand Orgue, the last two bandwidths were smaller than the preceding ones. This choice was made in order to better match the number of visible pipes in this region. An alternate choice could have been a doubled 1 kHz bandwidth only, essentially repeating the same values for the two last pipes. Nevertheless, due to the curved shape of the organ, the smallest pipes were often partly hidden by larger ones, and this issue was not too critical in the current installation.

Calibration. The third and most difficult part was the calibration of the VU-meter activations through a scaling of the frequency band dynamics to values ranging from 0 to 1. The null value corresponds to an empty VU-meter (no sound energy in this frequency band), and 1 to a full

Table 1. Frequency bandwidths

<i>Récit and Positif Bandwidth (33 bands)</i>	<i>Central Bandwidth (35 bands)</i>	<i>Lateral Bandwidth (20 bands)</i>
0 - 600 Hz 5 x 120 Hz	0 - 1200 Hz 10 x 120 Hz	0 - 600 Hz 5 x 120 Hz
600 - 1600 Hz 5 x 200 Hz	1200 – 3200 Hz 10 x 200 Hz	600 - 2200 Hz 5 x 320 Hz
1600 - 5000 Hz 10 x 340 Hz	3200 - 6400 Hz 8 x 400 Hz	2200 - 5000 Hz 5 x 360Hz
5000 - 8000 Hz 5 x 600 Hz	6400 - 11000 Hz 5 x 900 Hz	5000 – 10000 Hz 5 x 1000 Hz
8000 - 12000 Hz 5 x 800 Hz	11000 – 12000 Hz 2 x 500 Hz	
12000 - 16000 Hz 3 x 1000 Hz		

Figure 5: Mapping between sound analysis and graphics



VU-meter (maximum overall amplitude for this frequency band). The sound output was calibrated by applying pre-calculated decibel shifts computed from spectral analysis of the preliminary recordings, so that 0 would correspond to the ambient sound with the air blower turned on. According to this analysis, each frequency band had approximately a 30 dB amplitude dynamic, therefore each VU-meter activation was divided by 30 so the VU-meter graphical rendering on the organ pipe would use the whole range of height positions during the entire course of the concert.

This technique for sound spectral analysis and calibration encountered the following difficulties:

1. The positions of the microphones varied slightly between each rehearsal and performance. Since the microphones could be close to different pipes depending on their positions in the organ divisions, it resulted in slight changes in the amplitude levels of the sound spectral analysis.
2. The acoustics of the church produced a slight feedback effect between microphones and loudspeakers, and the offsets of the VU-meter calibration had to be readjusted for each concert.

3. Since the dynamics of the pipes depended on the loudness of each concert piece, and since the concert pieces varied from very loud to quiet ones, these variations resulted either in saturation or in a lack of reaction of the VU-meters.
4. The electric bellows system for air pressure generated a low-frequency parasite noise that was captured by the microphones and had to be taken into consideration for the minimal calibration level.
5. Even though the microphones were placed inside the organ divisions, the sound of the instrument could interfere with the church sounds such as audience applause and loudspeakers. Some of the spectators noticed the influence of their hand clapping on the virtual VU-meter levels, eventually using this unintended mapping to transform the instrument into an applause meter.

Because of these difficulties, approximately half an hour before the beginning of each concert was devoted to manual correction of the pipe calibration. The lowest activation level of the pipes was tuned with the bellows system switched on to cope with this background noise. In order to deal with the variations of dynamics between the concert pieces, the dynamics of each organ division was controlled by a slider on the audio monitoring interface. Last, the applause effects were avoided by manually shifting down all the division sliders after each piece.

Broadcast. The third and last module task was to concatenate all the frequency band values into a single message that could be sent through UDP over the Ethernet network. All the values were scaled to an numerical range [0, 1]. The real-time intensity values were doubled and a memory of the last maximal value was kept so that the audience would get the impression of a real VU-meter with an instantaneous value and a peak-hold maximum value. If no instantaneous value would overpass the peak-hold value for a half second, the current intensity value would replace the last peak-hold value. Thus two lists of 141 values were sent to the 3D graphical engine through UDP messages over the internal Ethernet network: 141 instantaneous frequency bands amplitudes and 141 associated last maximal values.

Audio Rendering

Algorithms were designed such that separate divisions of the organ were processed separately as they have different tonal properties (timbre, dynamics, and pitch range) and often contrast each other. Microphone signals were digitally converted via multi-channel audio cards with low-latency drivers; the real-time audio processing was implemented in Pure Data (Puckette, 1996). Selected algorithms included ring modulation, harmonizers, phasers, string resonators, and granular synthesis. Audio rendering was reproduced over an 8-channel full-bandwidth speaker configuration along the perimeter of the audience area and an additional high-powered subwoofer on the altar of the church, at the end opposite from the organ.

Historically, the rich variety of organ sounds is based on the combination of pipe ranks or registers. Various audio effects can be added as “electronic registers” to pipe organs, but their musical practicability strongly depends on the acoustics of the different pipes. Flue pipes, for example, have a frequency spectrum which is sparse and limited to a few harmonics for some stop ranks, e.g. the *Bourdon*; and as such are

well suited to additive synthesis algorithms creating more harmonically rich sounds. On the contrary, reed pipes offer a very dense frequency spectrum with high dynamics that might “overload” additive audio effects, which could yield distortions and noise-like sounds. However, subtractive synthesis algorithms are well suited to reed pipes as they allow one to spectrally reshape the harmonically rich organ sounds.

Ring modulators and harmonizers fall into the category of additive effects and have been well studied in signal processing literature (Zölzer, 2002; Verfaillie, 2006). Ring modulation, or double sideband (DSB) modulation, can be realized by multiplying two signals together producing components equal to two times the number of frequency components in one signal multiplied by the number of frequency components in the other. Harmonizing relates to adding several pitch-shifted versions of a sound to itself and various shifting ratios are used in order to produce different degrees of inharmonicity. In practice, microphones capture the global sound of each organ division, providing a polyphonic input signal to the harmonizers. The many inharmonic partials are added to the original signal spectrum to produce a very dense and inharmonic sound. The Karplus-Strong string resonator is a physical model-based algorithm simulating plucked string sounds, and is closely related to digital waveguide sound synthesis (Karjalainen et al., 1998) which provides a computationally efficient and simple approach to subtractive synthesis (Karplus & Strong, 1983). The algorithm consists of variable delay lines and low-pass filters arranged in a closed loop which allow dynamic control of resonance effects. When applied to the rich spectrum of reed pipes, this algorithm results in human voice like sounds with rapidly changing formants.

A large variety of multi-channel spatial audio systems have been developed recently such as quadrophony, vector base amplitude panning (VBAP), wave field synthesis (WFS), and Ambisonics. The sound spatialization environment used for the ORA project relied on third-order Ambisonics for 2D sound projection

on the horizontal plane only. Ambisonics was invented by Gerzon (1973). While the church room acoustic generates reverberation that is part of the organ sound, the presence of early reflections and late reverberation deteriorates sound localization accuracy. Different weighting functions described in (Noisternig et al., 2003) were applied prior to the Ambisonic decoding in order to widen or narrow the directional response patterns. The ORA sound design was made tolerant to reduced localization accuracy by employing non-spatially focused sounds, variable room reverberation algorithms, and spatial granular synthesis.

The classical repertoire organ pieces were spatialized and the reverberation time of the church acoustics was digitally increased. For these pieces, it made the organ sound independent from the organ location and it rendered the room sound much larger than the actual Sainte Elisabeth church. For the contemporary piece with audio digital effects, the captured sounds were distorted in real-time through signal processing algorithms.

CONCLUSION AND PERSPECTIVES

The audio and graphical calibrations of the installation involved manual adjustments that could be avoided through automatic calibration equipments and algorithms. By equipping the organ facade with fiducials (graphical patterns that can be captured by video camera and recognized by visual pattern-matching algorithms), the VU-meter quads could be automatically registered with the organ pipes. Such a realignment would make the installation more robust to slight displacements of the video-projectors between concerts and/or rehearsals. On the audio calibration side, background noise detection could be improved by automatically capturing the decibel amplitudes of the background noise frequency bands and by using these values to calibrate the minimal intensity values of the VU-meters. Similarly, automatic detection of spectral band maximal values could be used for amplitude calibration so that each VU-meter

could use the full range of graphical heights during the concert.

The ORA project has shown that the audience is receptive to such a new mode of instrument augmentation that does not burden artistic expression with additional and unnecessary complexity, but instead subtly reveals hidden data, making the performance both more appealing and better understandable. The work presented in this article opens new perspectives in musical instrument augmentation. First, graphical ornamentation could be applied to smaller non-static musical instruments such as string instruments by tracking their spatial location. Second, digital graphics for information visualization could reveal other hidden physical data such as air pressure, keystrokes, or valve closings and openings. This could be made possible by equipping the instrument with other additional sensors in addition to the microphones. Information visualization could also deal with a fine capture of ambient sounds such as audience noise, acoustic reflections, or even external sound sources such as street noise, and use them as additional artistic elements.

In summary, this project has demonstrated that live electronics applied to the pipe organ can extend the musical capacities and repertoire of the instrument, while maintaining its historical character. As such, it is then possible to mix classical and contemporary music harmoniously. The ORA augmented instrument thus offers performers and composers new means of expression.

REFERENCES

- Bouillot, N., Wozniowski, M., Settel, Z., & Cooperstock, J. R. (2007). A mobile wireless augmented guitar. In *Proceedings of the 7th International Conference on New Interfaces for Musical Expression NIME '07*, Genova, Italy.
- Cakmakci, O., Bérard, F., & Coutaz, J. (2003). An augmented reality based learning assistant for electric bass guitar. In *Proceedings of the 10th International Conference on Human-Computer Interaction (HCI International 2003)*, Crete, Greece.

- d'Alessandro, C., Noisternig, M., Le Beux, S., Katz, B., Picinali, L., Jacquemin, C., et al. (2009). The ORA project: Audio-visual live electronics and the pipe organ. In *Proceedings of International Computer Music Conference ICMC 2009*, Montreal, Canada.
- Fels, S., Nishimoto, K., & Mase, K. (1998). Musikalscope: A graphical musical instrument. *IEEE MultiMedia*, 5(3), 26–35. doi:10.1109/93.713302
- Gerzon, M. A. (1973). Periphony: With-height sound reproduction. *Journal of the Audio Engineering Society. Audio Engineering Society*, 21(1), 2–10.
- Jacquemin, C., & Gagneré, G. (2007). Revisiting the Layer/Mask Paradigm for Augmented Scenery. *International Journal of Performance Arts and Digital Media*, 2(3), 237–258. doi:10.1386/padm.2.3.237_1
- Jacquemin, C., Planes, B., & Ajaj, R. (2007). Shadow casting for soft and engaging immersion in augmented virtuality artworks. In *Proceedings of 9th ACM Conference on Multimedia 2007*, Augsburg, Germany.
- Jordà, S. (2003). Interactive music systems for everyone: Exploring visual feedback as a way for creating more intuitive, efficient and learnable instruments. In *Proceedings of the Stockholm Music Acoustics Conference (SMAC 03)*, Stockholm, Sweden.
- Jordà, S., Geiger, G., Alonso, M., & Kaltenbrunner, M. (2007). The ReacTable: exploring the synergy between live music performance and tabletop tangible interfaces. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction TEI '07* (pp. 139-146). New York: ACM.
- Karjalainen, M., Valimi, V., & Tolonen, T. (1998). Plucked String Models: From the Karplus-Strong Algorithm to Digital Waveguides and Beyond. *Computer Music Journal*, 22(3), 17–32. doi:10.2307/3681155
- Karplus, K., & Strong, A. (1983). Digital Synthesis of Plucked String and Drum Timbres. *Computer Music Journal*, 7(2), 43–55. doi:10.2307/3680062
- Levin, G., & Lieberman, Z. (2004). In-situ speech visualization in real-time interactive installation and performance. In *Proceedings of the 3rd international symposium on Non-photorealistic animation and rendering NPAR '04* (pp. 7-14). New York: ACM.
- Lewis Charles Hill, I. (2006). *Synesthetic Music Experience Communicator*. Unpublished doctoral dissertation, Iowa State University, Ames, IA.
- Marrin, T., & Paradiso, J. (1997). The Digital Baton: a Versatile Performance Instrument. In *Proceedings of the International Computer Music Conference ICMC 1997*.
- Miranda, E. R., & Wanderley, M. (2006). *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard (Computer Music and Digital Audio Series)*. Madison, WI: A-R Editions, Inc.
- Motokawa, Y., & Saito, H. (2006, October 22-25). Support system for guitar playing using augmented reality display. In *Proceedings of the 2006 Fifth IEEE and ACM International Symposium on Mixed and Augmented Reality (Ismar '06)* (pp. 243-244). Washington, DC: IEEE Computer Society.
- Noisternig, M., Sontacchi, A., Musil, T., & Höldrich, R. (2003). A 3D ambisonic based binaural sound reproduction system. In *Proceedings Audio Engineering Society AES 24th International Conference*, Banff, Canada.
- Poupyrev, I., Berry, R., Billinghamurst, M., Kato, H., Nakao, K., Baldwin, L., et al. (2001). Augmented Reality Interface for Electronic Music Performance. In *Proceedings of HCI 2001* (pp. 805-808).
- Puckette, M. S. (1996). Pure data: Another integrated computer music environment. In *Proceedings of the International Computer Music Conference ICMC 1996*, Hong Kong, China (pp. 37-41).
- Raskar, R., & Beardsley, P. (2001). A self-correcting projector. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2001*, Kauai, HI (pp. 504-508). Washington, DC: IEEE Computer Society.
- Rocha, F., & Malloch, J. (2009). The Hyper-Kalimba: Developing an Augmented Instrument from a Performer's Perspective. In *Proceedings of the 6th Sound and Music Computing Conference SMC 2009*, Porto, Portugal (pp. 25-29).
- Thompson, J., & Overholt, D. (2007). Sonofusion: Development of a multimedia composition for the overtone violin. In *Proceedings of the International Computer Music Conference ICMC 2007 International Computer Music Conference*, Copenhagen, Denmark (Vol. 2).
- Verfaillie, V. (2006). Adaptive Digital Audio Effects (A-DAFx): A new class of sound transformations. *IEEE Transactions on Audio, Speech and Language Proc.*, 14(5), 1817–1831. doi:10.1109/TSA.2005.858531
- Welch, P. (1967). The use of Fast Fourier Transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics, AU-15*, 70–73. doi:10.1109/TAU.1967.1161901
- Zölzer, U. (2002). *DAFx – Digital Audio Effects*. New York: John Wiley and Sons.

ENDNOTES

- ¹ <http://toddvanderlin.com/2009/04/ar-scratching/> <http://www.synthtopia.com/content/2009/04/28/augmented-reality-dj-scratching/>
- ² The ORA project participants were: Organ performer: Christophe d'Alessandro ; Live electronics performer: Markus Noisternig ; Video installation: Bertrand Planes ; Audio design: Christophe d'Alessandro, Markus Noisternig, Sylvain Le Beux, Lorenzo Piccinali, Brian Katz, Nicolas Sturm, Nathalie Delprat; Video design: Rami Ajaj and Christian Jacquemin. The Musical Works were compositions of F. Couperin, J.S Bach, C. Franck,

O. Messiaen, C. d'Alessandro. Videos of the event can be found at <http://www.youtube.com/watch?gl=FR&hl=fr&v=JIYVUtJsQRk> or in higher resolution at the project site http://vida.limsi.fr/index.php/Orgue_Augmentee. http://vida.limsi.fr/index.php/Realite_Augmentee_Mobile

³ In scientific articles, augmented instruments are often called hybrid instruments, hyperinstruments, or extended instruments.

⁴ <http://virchor.sf.net>

⁵ Inkscape is a vector graphic editor: <http://www.inkscape.org>

⁶ Since only few pipes are visible, the exact note of a pipe was not necessarily falling into the frequency band displayed on its surface.

⁷