Networked Head-Mounted Displays for Animated Notation and Audio-Scores with SmartVox

Jonathan Bell
Aix Marseille Univ, CNRS, PRISM
Perception, Representations, Image, Sound, Music
Marseille, France
belljonathan50@gmail.com

ABSTRACT

The combination of graphic/animated scores, acoustic signals (audio-scores) and Head-Mounted Display (HMD) technology offers promising potentials in the context of *distributed notation*, for live performances and concerts involving voices, instruments and electronics. After an explanation of what SmartVox is technically, and how it is used by composers and performers, this paper explains why this form of technology-aided performance might help musicians for (spectral) tuning and synchronization with an electronic tape. Then, from an exploration of the concepts of *distributed notation* and *networked music performances*, it proposes solutions (in conjunction with INScore, BabelScores and the Decibel Score Player) seeking for the expansion of distributed notation practice to a wider community. It finally presents findings relative to the use of SmartVox with HMDs.

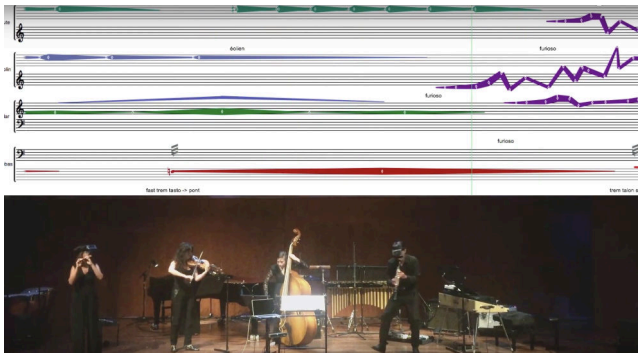

Figure 1: Instrumentalists using SmartVox with head-mounted displays at Gaudeamus Festival.

Author Keywords

Networked Music Performance, Computer-aided performance, Computer-aided composition, Musical notation, Head-Mounted Display, Augmented Reality

CCS Concepts

•**Networks** → **Local area networks;** •**Human-centered computing** → *Smartphones;* •**Applied computing** → *Computer-assisted instruction; Performing arts;* **Sound and music computing;**

NIME'19, June 3-6, 2019, Federal University of Rio Grande do Sul, Porto Alegre, Brazil.

1. INTRODUCTION: WHAT IS SMARTVOX?

SmartVox [4][3] is a distributed web application that delivers and synchronizes audiovisual scores in the mp4 format to the performer's mobile devices, in compositions involving up to 80 simultaneous performers such as in ***Le temps des Nuages***. Recent developments include the use of head-mounted displays (HMDs) for technology-aided performance, as in the pieces ***In Memoriam Jean-Claude Risset I*** (Figure 1) and ***Mit Allen Augen, In Memoriam J.C. Risset II*** (Figure 2).[1]

Figure 2: Singer wearing HMD for technology-aided performance.

1.1 Technical implementation

SmartVox was developed in the *SoundWorks* framework.[2] *SoundWorks* provides a set of services – such as synchronization, network messages, distributed states, creation of groups of clients – that aims to solve problems common to distributed and synchronized web applications centered on multimedia rendering. The framework is written in Javascript, with a server side based on Node.js. [3]The SmartVox application consists of two web clients, the player and the conductor, that can be executed in any recent web browser on mobile devices (e.g. smartphones, tablets) and laptops. The real-time communication between clients is achieved through the WebSocket protocol [4] The application is typically deployed

[1]Those three pieces are respectively available at: https://youtu.be/SyFdR2HiF00, https://youtu.be/hQtyu1dcCaI, and https://youtu.be/ET_OBgFWx04.
[2]*SoundWorks* was initiated by the CoSiMa research project funded by the French National Research Agency (ANR) and coordinated by Ircam.
[3]https://nodejs.org/en
[4]https://www.w3.org/TR/WebSockets/

over a *local area network*, but it may also be used over the internet.[5]

1.2 Composer's Score Elaboration Workflow

The audiovisual scores (mp4 files) used in Smartvox are composed in the Bach environment [1], a Max library for real-time computer-aided composition. In this environment (bach.roll or bach.score), each note of the score can store metadata in its *slot content* (e.g. Path to a sample, automation curve...). This feature is used to control synthesizer that creates the audio cues, directly from the notational environment [6]. Each separate part (one part per voice/instrument) of audiovisual notational material generated in Bach is then recorded by screen/audio captures, in order to be distributed and synchronized by SmartVox in rehearsals and performance, as shown in the diagram below:

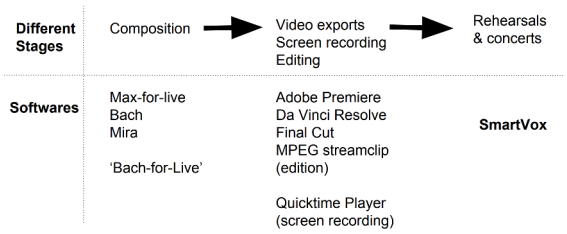

Figure 3: Score elaboration workflow.

1.3 Performer's Perception

From a singer's or instrumentalist's point of view, SmartVox provides simultaneously auditory (audio-score[2]) and visual (animated notation) information, which facilitates the interpretive and performative processes by recalling of the performers' separate parts, thus expediting the learning and rehearsal process of a new work. This synergy between visual and acoustic stimuli also enables spatial separation or movement of the performers, and allows for the exploration of unfamiliar musical materials (e.g. microtonal tuning, texts in a foreign language).

2. SMARTVOX FOR MIMETIC ORCHESTRATION OF ELECTRONICS

2.1 Frequency Domain : Spectral Harmony

SmartVox was initially imagined as a tool that would help choristers to tune with the electronics. In spite of its initially loose synchronization capabilities, it quickly appeared that, with singers of all levels, the system allowed for a precise control of harmony in the context of choral music in particular.[7]The piece ***Smartvox***[8] (which has the same name as the application) remains the best illustration of it capabilities in spectral terms.

2.2 Time Domain: Rhythmic Precision

At IRCAM in January 2018, Benjamin Matuszewski improved the synchronization possibilities of SmartVox by implementing a client-side algorithm that puts back in sync devices whose drift exceeds a certain threshold ([3], Chapter 4.3). Thanks to this new implementation, it can now explore a wide range of tightly synchronized musical situations. These new possibilities encouraged the author to write rhythmical games between performers, taking advantage of the cursor-type display of time, as in ***In Memoriam Jean-Claude Risset 1***,[9] or sharper articulations, such as the *clapping* game present throughout the piece ***Mit Allen Augen, in memoriam J.C. Risset II***.[10]

During the composition process,[11] the precise lining-up and synchronization of live performers with electronics was also improved by ability to use *Bach* as a *master* over *Ableton* : the bach.roll object displays notation in proportional time, and outputs notifications of its playback status in real-time. These notifications can be interpreted in *Max For Live* in order to synchronize 1/ the notation for human players in *Bach* and 2/ the electronics in Ableton Live[12].

2.3 Conclusion : Mimetic Orchestration

A great majority of the electronic material presented in the *Risset cycle*[13] constitutes *mimetic* [14]orchestrations of the PRISM laboratory synthesizer, developed by Richard Kronland-Martinet [6]. The recent improvement of SmartVox in terms of synchronisation now allow more flexibility for musical expermentation of various sorts. In ***Das Hoheslied, In Memoriam Jean-Claude Risset 3*** for instance, students of HfMT Hamburg were able to sight-sing (for the very first time) passages which strikingly resembled the target model.[15]

3. DISTRIBUTED NOTATION

3.1 Distributed Notation: Global Context

Composers and researchers increasingly acknowledge the strong analogy which can be drawn between the traditional 'score and parts' musical practice led by a conductor, and the modern distributed systems or web applications (MASD [15] - SmartVox [4]), in which multiple clients coordinate their actions by passing messages to one another. Consequently many performance-oriented systems today (INScore[7], SmartVox [4][3], Decibel [14] [13], Maxscore [12]) endeavor to distribute and synchronize each part of the score on the performer's devices (whether Smartphones, tablets or laptops). The Tenor 2018 conference in Montreal revealed to the author an interesting similarity between the Decibel ScorePlayer and SmartVox, in the sense that the two softwares are rather elementary solutions both converging towards the *score player* performance-oriented paradigm, whilst other notation packages (Bach, Maxscore, Symbolist...) were also designed for more elaborate computer-aided composition and real-time processes. SmartVox reveals itself to be very effective for this kind of setup by rendering the score in the browser directly, without the need for installing a client application. Thanks to cross-platform web technologies, the application works with any browser capable

[5]https://youtu.be/83ub6-Q5oj0
[6]See https://youtu.be/s4qS2khwkT0 for demonstration.
[7]See https://youtu.be/rlmqnIJhayU for demonstration.
[8]See https://youtu.be/JZsJn7EEW-A.
[9]A rhythmical passage in *In Memoriam Jean-Claude Risset 1* https://youtu.be/hQtyu1dcCaI?t=349
[10]See https://youtu.be/ET_OBgFWx04?t=618.
[11]The compositional stage is made in *Bach* and *Ableton*, as opposed to the performance during which only *SmartVox* is running
[12]The following example shows the convenience of the Bach/Ableton inter-application communication https://youtu.be/VJvY5wYl_cM
[13]*In Memoriam J.C. Risset*, *Mit Allen Augen, In Memoriam J.C. Risset II*, *Das Hoheslied, In Memoriam Jean-Claude Risset III*
[14]A term attributed to Lindsay Vickery [16] and Benedict Carey [5].
[15]A *Target sound* is followed by its imitation by the choir in a sight-reading session of *Das Hoheslied, In Memoriam Jean-Claude Risset 3*. See https://youtu.be/EHYq9nFF6sE for demonstration.

device, and no installation is required by the client. The *node.js*[16]/WebSocket architecture of SmartVox will hopefully inspire more composers, researchers and developers to investigate the emerging musical practice of distributed notation.

3.2 SmartVox.eu, Score Distribution over the Web and Remote NMPs

With the exponential growth of the web, hosting ***SmartVox***[17] on the internet (i.e. on a remote server rather than a local one) appeared as a necessity. However we demonstrated (in Chapter 4.2 'Measurements of timing accuracy' [3]) that, although the synchronization of different parts was quite accurate in this way, local solutions remained safer (in the same room, over Wi-Fi). Also, notwithstanding the fact that SmartVox belongs to the realm of Networked Music Performances, it became clear that its practical application falls under the 'local NMPs' sub-category defined by Gabrielli and Squartini [10]. Indeed in the case of SmartVox today, local NMPs still seem more viable, first technically (because of the unreliability of the internet in 2019), but also artistically speaking (because in a live performance the audience presumably expects the physical presence of the artist [11]).

3.3 The 'BabelBox', a 0-conf Raspberry Pi Solution for Local NMPs

The author has shown evidence that SmartVox is suitable for large scale projects such as ***le temps des nuages***[18] in which eighty singers and musicians had their score synchronized through the same network [3]. However, in the more intimate context of chamber music (as in ***In Memoriam JC Risset 1***), the focus is slightly different: the network load is much lighter, which encouraged finding a minimal hardware solution, in order to make possible rehearsals without the physical presence of the composer or the mobilization of a technician only to setup a network. Installing the server on the performer's computer remotely has often been successful even with musicians unfamiliar with technology, thanks to the flexibility of the Node Package Manager (NPM), which reduces the installation of SmartVox to a few command lines.[19]This installation process nevertheless remained an obstacle for the dissemination of SmartVox. In search of a light plug-and-play dedicated system to be sent over the post, the Raspberry Pi quickly appeared as the best option to host SmartVox on an embedded system. *node.js* runs on Raspbian, and SmartVox proved to be very stable on a Raspberry Pi 3, so, once installed, the only two steps for a *0-conf* deliverable hardware were:

- Setting up a static address for a dedicated router (e.g. tp-link...).
- Starting SmartVox at boot.

This low-cost system (less than 65 €, for a Raspberry and a router) now allows the sending of ready-to-use scores. Once the system is power-supplied, all the performers need to do is to join the dedicated Wi-Fi, and type the static IP address of the server on their smartphone/tablet (i.e. for the performers: 192.168.0.100:8000, and for the conductor: 192.168.0.100:8000/conductor). In January 2019, for the performance of a pedagogical piece, ***The Super Mario Babel Box***. This small system was sent per post to the Caen French conservatoire via BabelScores[20], thus proposing a rental of performing scores (separate parts) of a new kind. During the rehearsals and performances, synchronization prooved to be of great precision and download speed very fast, in spite of the lightness of the low-cost system.

3.4 BabelScores, Inscore and SmartVox

The representation of time in the age of screen scores is open to a wide array of solutions. In none-pulsed music particularly, and in spite of the great conducting tradition in chamber music and orchestral works, left-to-right scrolling cursors distributed on the performers' devices (as in the Decibel ScorePlayer or SmartVox) seem a far more straightforward strategy to obtain tight synchronization in comparison to the bars and beats 'encoding' (quantification by the composer with or without the help of algorithms) and decoding processes (a compromised interpretation by the instrumentalist, between the rhythmic values written on the page and the gestures of the conductor), inherited from a scoring tradition, in which a regular meter was assumed.

If *Bach* is the favoured environment for the author's own compositions, INScore[9][8] might be of particular relevance to extend the practice of distributed notation to a wider community of composers and performers, through the wide community of composers edited edited at BabelScores in particular. INScore allows for the precise temporal control of animated cursors in the graphical domain. Thanks to its OSC support, it is possible to control the cursor's position via automations in Ableton Live[21]. These animation can be used for scrolling cursors over standard notation[22], and the arm movements of a conductor can be replaced by a bouncing ball[23]. These tools might be used, in combination with the BabelBox, for highly demanding pieces in terms of synchronization[24], or as a tool for appreticeship of unfamiliar idioms such as early music notation.[25]

4. HEAD MOUNTED DISPLAYS

In Memoriam JC Risset 1[26], premiered in September 2018 at the Gaudeamus Festival (Utrecht), constitutes the author's first experiment using head-mounted displays. By simply displaying each part of the score over the heads of the performers (for flute and clarinet only), this piece revealed interesting potentials, further exploited on a larger scale in ***Mit Allen Augen, In Memoriam J.C. Risset II***.

Just as traditional scores placed on a music stand, screen-scores displayed on a tablet (for instrumentalists) or on a phone (for singers) oblige musicians to look and orientate their body constantly in the direction of the score. This well-established convention of the classical concert setup considerably limits the possibilities of staging music, in a theatrical context for instance.

According to the performers' feedback, head-mounted setups provide a large and comfortable display, since the

[16]Server-side JavaScript, see: https://nodejs.org/en/

[17]Each instrumental/vocal part of the piece *And the Sea* is accessible through the following url www.smartvox.eu, and can be accessed simultaneously from e.g., an iPad for the flute, Android tablets for piano and cello, and a phone for the singer. A trailer of the premiere of the piece is accessible here: https://youtu.be/prcXUbhd-ZY

[18]A recording of the piece is available at the following address: https://youtu.be/SyFdR2HiF00

[19]SmartVox is open source and ready to download on GitHub: https://github.com/belljonathan50/SmartVox0.1.

[20]Babelscores (https://www.babelscores.com/) is an online score database for classical contemporary music, currently actively supporting the SmartVox project: http://1uh2.mj.am/nl2/1uh2/lgi4u.html.

[21]See https://youtu.be/rLy8DW_p2JE for demonstration.

[22]See https://youtu.be/y9oSNATeVXk.

[23]See https://youtu.be/isbcStnt0_k.

[24]Seehttps://youtu.be/0y54fwJSFmY?t=10.

[25]Seehttps://youtu.be/DaMAnO040NI.

[26]A recording of the piece is available at the following address: https://youtu.be/hQtyu1dcCaI

environment is still visible around the score, or through the score in the case of holographic display (such as the Aryzon headset for instance, see Figure 4). The performers also showed interest in their ability to move freely on stage oraround the audience.

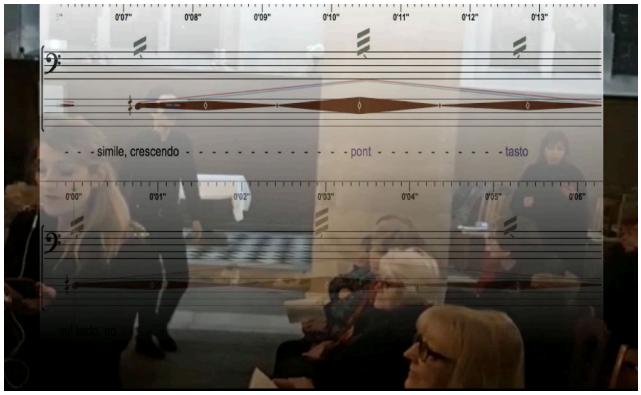

Figure 4: Score display from the performer's point of view.

While awaiting more discrete solutions (such as *Vufine*[27] Glasses), invisible from an audience's point of view, HDMs open vast perspectives for further research, which could expand even beyond the mere display of animated notation.

5. CONCLUSION

Initially conceived as a rehearsal tool for choral practices almost exclusively [4] [3], recent use cases have demonstrated that SmartVox is in fact better and more broadly described as a *distributed score player* suitable to instruments as well as vocalists with, or without electronics. If musical notation remains a key concept for SmartVox, the notion of *Networked Musical Performance* probably defines it more specifically. With its promising use of distributed head-mounted displays, SmartVox wishes to shift the focus from human-computer interaction towards computer-mediated human interaction and proposes an original interface for musical expression for collective technology-aided live performances.

6. ACKNOWLEDGMENTS

Andrea Agostini, Richard Baker, Laurence Brisset, Dominique Fober, Daniele Ghisi, Georg Hajdu, Richard Kronland-Martinet, Benjamin Matuszewski and Norbert Schnell.

7. REFERENCES

[1] A. Agostini and D. Ghisi. Bach: an environment for computer-aided composition in max. In *Proceedings of the 38th International Computer Music Conference (ICMC)*, Ljubljana, Slovenia, 2012.

[2] J. Bell. *Audio-scores, a resource for composition and computer-aided performance*. PhD thesis, Guildhall School of Music and Drama, 2016.

[3] J. Bell. Audivisual Scores and Parts Synchronised over the Web. In *TENOR 2018*, Montreal, France, 2018.

[4] J. Bell and B. Matuszewski. SmartVox. A web-based distributed media player as notation tool for choral practices. In *Proceedings of the 3rd International Conference on Technologies for Music Notation and Representation (TENOR)*, Coruña, Spain, 2017. Universidade da Coruña.

[27] https://store.vufine.com/

[5] B. E. Carey. *From Mimetics to Memetics*. PhD thesis, Hochschule für Musik und Theater Hamburg, (unpublished).

[6] S. Conan, E. Thoret, M. Aramaki, O. Derrien, C. Gondre, S. Ystad, and R. Kronland-Martinet. An intuitive synthesizer of continuous-interaction sounds: Rubbing, scratching, and rolling. *Computer Music Journal*, 38(4):24–37, Dec 2014.

[7] D. Fober, G. Gouilloux, Y. Orlarey, and S. Letz. Distributing music scores to mobile platforms and to the internet using inscore. In *Proceedings of the Sound and Music Computing conference — SMC'15*, pages 229–233, 2015.

[8] D. Fober, Y. Orlarey, and S. Letz. Inscore time model. In *Proceedings of the International Computer Music Conference*, pages 64–68, 2017.

[9] D. Fober, Y. Orlarey, and S. Letz. Towards dynamic and animated music notation using inscore. In V. Ciciliato, Y. Orlarey, and L. Pottier, editors, *Proceedings of the Linux Audio Conference — LAC 2017*, pages 43–51, Saint Etienne, 2017. CIEREC.

[10] L. Gabrielli and S. Squartini. *Wireless Networked Music Performance*, pages 53–92. Springer Singapore, Singapore, 2016.

[11] G. Hajdu. Embodiment and disembodiment in networked music performance. In *Body, Sound and Space in Music and Beyond: Multimodal Explorations*. Taylor & Francis, 2017.

[12] G. Hajdu and N. Didkovsky. Maxscore: Recent developments. In S. Bhagwati and J. Bresson, editors, *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR'18*, pages 138–146, Montreal, Canada, 2018. Concordia University.

[13] C. Hope, L. Vickery, A. Wyatt, and S. James. The decibel scoreplayer - a digital tool for reading graphic notation. In M. Battier, J. Bresson, P. Couprie, C. Davy-Rigaux, D. Fober, Y. Geslin, H. Genevois, F. Picard, and A. Tacaille, editors, *Proceedings of the First International Conference on Technologies for Music Notation and Representation – TENOR'15*, pages 58–69, Paris, France, 2015.

[14] C. Hope, A. Wyatt, and D. Thorpe. Scoring an animated notation opera – the decibel score player and the role of the digital copyist in 'speechless'. In S. Bhagwati and J. Bresson, editors, *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR'18*, pages 193–200, Montreal, Canada, 2018. Concordia University.

[15] D. G. Nathan Magnus. Musician assistance and score distribution (masd). In *Proceedings of The International Conference on New Interfaces for Musical Expression – NIME'2012*, Ann Arbor., 2012. University of Michigan.

[16] L. Vickery. Hybrid real/mimetic sound works. In R. Hoadley, C. Nash, and D. Fober, editors, *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR'16*, pages 19–24, Cambridge, UK, 2016. Anglia Ruskin University.