PARTIELS - EXPLORING, ANALYZING AND UNDERSTANDING SOUNDS

Pierre Guillot

IRCAM - Centre Pompidou
STMS UMR 9912
Paris, France
pierre.guillot@ircam.fr

ABSTRACT

This article presents Partiels, an open-source application developed at IRCAM to analyze digital audio files and explore sound characteristics. The application uses Vamp plug-ins to extract various information on different aspects of the sound, such as spectrum, partials, pitch, tempo, text, and chords. Partiels is the successor to AudioSculpt, offering a modern, flexible interface for visualizing, editing, and exporting analysis results, addressing a wide range of issues from musicological practice to sound creation and signal processing research. The article describes Partiels' key features, including analysis organization, audio file management, results visualization and editing, as well as data export and sharing options, and its interoperability with other software such as Max and Pure Data. In addition, it highlights the numerous analysis plug-ins developed at IRCAM, based in particular on machine learning models, as well as the IRCAM Vamp extension, which overcomes certain limitations of the original Vamp format.

1. INTRODUCTION

Partiels[1] is an application for analyzing digital audio files, designed to provide a dynamic, ergonomic interface for exploring the content and characteristics of sound (spectrum, fundamental frequency, tempo, chords, spectral centroid, speech analysis, etc.). The Partiels application is available for Windows, macOS, and Linux.

The application is based on the Vamp analysis plug-in format, developed at C4DM (Centre for Digital Music) - Queen Mary University of London. In parallel with the development of the main application, Vamp plug-ins were developed using IRCAM audio engines (notably SuperVP [1], PM2 [2][3], IrcamBeat [4] and IrcamDescriptors [5]) and external technologies (notably Whisper [6] and Crepe [7]). Moreover, Partiels lets you use all existing Vamp analysis plug-ins (such as those from the BBC, C4DM, Pompeu Fabra University, etc.)[2].

The Partiels application and its plug-ins have been developed at IRCAM within the Innovation and Research Means department since October 2020. The first version of Partiels was published on the IRCAM Forum in September 2021. In December 2023, the application and plug-ins became free, and since September 2024, Partiels has been available as an open-source project on GitHub under the GPLv3 license.

---

[1]https://github.com/Ircam-Partiels/Partiels
[2]https://www.vamp-plugins.org/

This article begins by presenting the background to the project and its objectives, and then goes on to describe the main features offered by the application[3]. Finally, it describes current developments and prospects for the project.

2. PROJECT ORIGINS

Partiels and the first analysis plug-ins associated with it derive from the AudioSculpt [8][9] application, a software package built around libraries and command-line tools, SuperVP, PM2, Ircam Descriptors, and Ircam Beat, developed at Ircam within the Sound Analysis and Synthesis team. AudioSculpt offered a graphical interface for displaying and interacting with analysis results and transforming sounds. The software is now deprecated, as its monolithic architecture and dependencies have prevented it from being updated and adapted to current hardware and operating systems[4]. The disappearance of this tool, popular within the IRCAM community, has created a major gap among signal processing researchers, musicologists, composers, computer music producers, teachers, sound designers, etc. It had become urgent to offer an alternative that would meet this demand, while avoiding the limitations that had led to the software's obsolete status.

The challenge was to offer a short-term solution to this problem. However, although they share similar practices and uses, the audiences concerned by this expectation have significantly different needs and uses. In signal processing, for example, it is important to be able to compare the results of different analysis engines. Musicologists, on the other hand, need to generate images and diagrams to analyze sounds. In composition, a frequent request is to be able to export analysis results for use in other software packages such as Max[5], Pure Data [10], and OpenMusic [11] to produce sounds or scores. Moreover, since the last version of AudioSculpt, new issues and practices have emerged, in line with the evolution of digital audio and computing in general.

---

[3]The application includes a large number of features that cannot be described in detail in this article, which aims to provide a general overview of the tool. Readers requiring further information on a specific feature can refer to the Partiels user manual.

[4]AudioSculpt is a mainly graphical application with a monolithic approach (all processing and analysis modules are embedded in the application). The application's code is based entirely on Apple's Carbon API, which is not cross-platform, has been deprecated since 2012, and is obsolete on 64-bit computers, meaning that it has not run on Apple computers for several years. As a result, the demand from users for an update to this software was growing. However, due to this monolithic approach and the obsolescence of almost all code, the development of a new version, even an elementary one, would require resources that are not available to IRCAM.
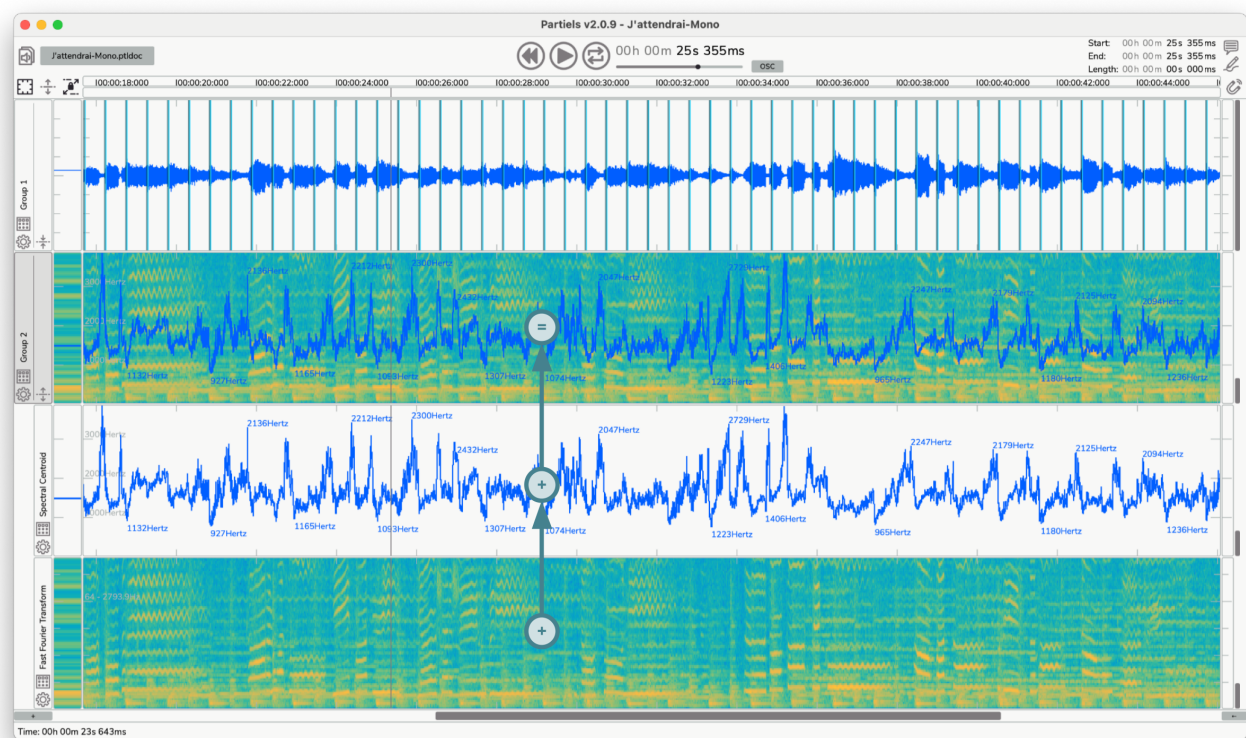
[5]https://cycling74.com/

Figure 1: *Overview of the Partiels application containing two groups of analyses. A first group (shrunk) with waveform and beat markers, and a second group (expanded) with spectrogram analysis and spectral centroid.*

The main functional specifications were as follows:

- to quickly and easily launch and modify analyses and their parameters,
- to correct analysis results textually and graphically,
- to improve visualization and comprehension of analyses,
- to facilitate communication and permeability with other software environments,
- to easily apply a set of analyses to a series of audio files,
- to compare several audio files and/or the different channels of the files,
- to share analyses and results with other users.

It was necessary to develop a tool powerful enough to respond to these complex and heterogeneous practices. The aim was to offer a relatively simple and ergonomic interface, to be accessible to a varied and potentially non-initiated public, while using a flexible and dynamic architecture that would allow it to evolve and integrate new technologies and analyses.

Other tools such as Praat [12], iAnalyse [13], Spek[6], and Sonic Visualizer [14][15] for sound analysis already exist, each offering solutions adapted to specific and varied needs. But none of these applications met all our specifications. One major obstacle was that, although Partiels is open-source, the algorithms developed by the Sound Analysis and Synthesis team are proprietary, preventing

their direct integration into other software. Sonic Visualizer, which is closest to our needs, has an architecture based on plugins, enabling the integration of external technologies. But the analyses developed at IRCAM have complex, multidimensional data models with different types and require dynamic inputs, which did not allow them to be integrated into this software. The new software solution proposed is based on the Vamp plugin format [16][14], incorporating an extension, Ircam Vamp Extension (IVE)[7], to meet specific needs.

3. GENERAL OVERVIEW

Partiels lets you control and apply analyses to digital audio files, organize, visualize, and edit results, and navigate through their representation. Data can be exported as images and in various text formats. Partiels offers additional features including a batch processing system, a command line interface, and the ability to send results via Open Sound Control (OSC) [17].

3.1. Organizing and managing analyses

A Partiels document is associated with one or more audio files. This document has one or more groups, which themselves include analysis tracks. These tracks are stacked one on top of the other

---

[6]https://www.spek.cc/

[7]https://github.com/Ircam-Partiels/ircam-vamp-extension

within a group. Groups are distributed one on top of the other within a document, as shown in Fig. 1.

Partiels is a plug-ins host. Without plug-ins, it can only perform waveform and spectrogram analyses. All complex analyses are performed by dynamically loaded Vamp plug-ins. At startup, the application searches for the plug-ins installed on the machine[8] and determines the various analyses available, which are then offered to the user. These analyses can generate three types of results, as established by the Vamp format, and which define their graphical representation mode: markers (time, duration, text) displayed as vertical lines to which text is potentially associated, points (time, duration, numerical value) displayed as segments, vectors (time, duration, list of numerical values) displayed as an image, most often for a spectrogram.

When an analysis is created, it is necessarily added to a group. It is then possible to modify the organization within a group and to move or duplicate an analysis from one group to another. It is also possible to modify analysis track parameters via the properties window. These parameters are divided into two categories, as shown in Fig. 2:
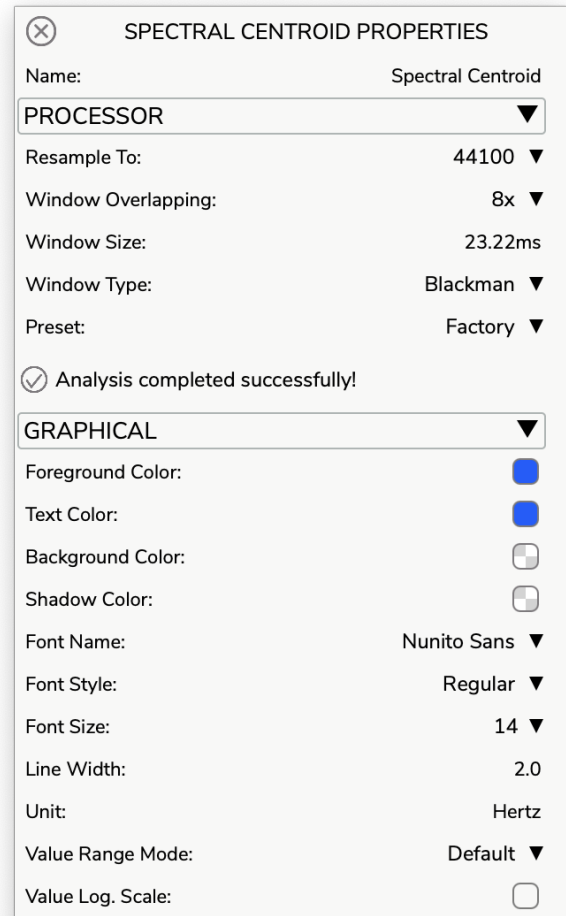

Figure 2: *Property window of a spectral centroid analysis track with a section corresponding to analysis engine parameters and a section corresponding to graphical rendering parameters.*

---

[8]The plug-in search parameters and access paths can be configured in the application properties

- Analysis engine parameters (such as Fourier transform window size, transient detection threshold, neural network model, etc.) that generate a new analysis on the fly when changed.

- Graphical rendering parameters (such as colors, font, text positioning, scale and range, grid, etc.) that generate on-the-fly updates of representations.

This approach enables graphical representations to be matched to the type of analysis and its context, in particular to highlight certain data in relation to others. On-the-fly updating of analyses and their representations is particularly useful for experimenting, understanding parameters, and adapting them as required.

3.2. Audio file playback and management

Partiels offers an engine with an audio file layout system, as shown in Fig. 3, and multichannel management linked to analysis. Each playback channel is associated with one channel of the audio file (or the sum of the channels).
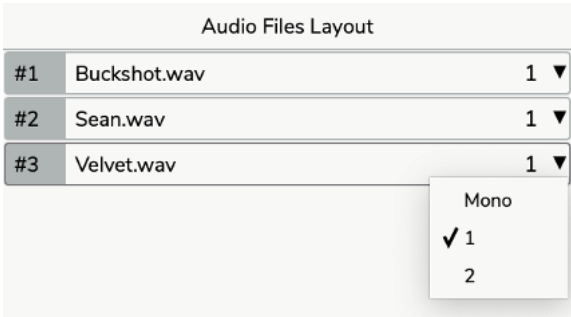

Figure 3: *Audio files layout window for analysis and playback with three files used for three separate channels.*

This approach makes it possible to compare different audio files and/or different channels of the same file in order, for example, to highlight differences in interpretations of the same piece, the artifacts of different encoding, or to enable analysis of audio spatialization effects (such as the spread of a stereophonic panning). During audio playback, each track in the layout system can be sent to a specific sound card channel via a routing matrix, in order to adapt to the analysis context and playback hardware. Finally, audio transport offers a looping system that can be magnetized to marker tracks to focus on specific time sections.

3.3. Visualizing analysis results

The main graphical interface represents the evolution of results over time, over which the playhead moves automatically during audio playback (or manually with the mouse).

3.3.1. Instant representation

The application also offers a representation of results on the left-hand side of the global interface, as well as a representation in the form of a table of numerical and textual values in a floating window, as shown in Fig. 4. These complementary representations are synchronized on the playhead, enabling joint reference to the different interfaces, so that the advantages of each type of visualization can be exploited simultaneously.

< **300** >

Figure 4: *Table of values containing the results of the marker values of an analysis with the VAX plug-in resulting from the realignment of the Whisper analysis, which generated the sentence "Bienvenue aux ateliers du Forum de l'IRCAM.".*

3.3.2. Open Sound Control

In addition, the results can be sent to third-party applications such as Pure Data, Max, or SuperCollider [18] via OSC, as shown in Fig. 5, to generate further representations (and/or sonify the results). The Partiels distribution provides example patches for this purpose.
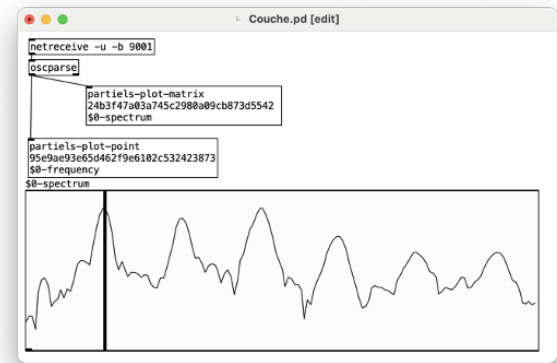

Figure 5: *Pure Data patch displaying the values of a spectrogram and a pitch detection based on data sent by Partiels in OSC.*

3.4. Navigating and editing of analysis results

Partiels offers two modes of interaction inspired by Max and Pure Data: navigation and editing. The navigation mode is dedicated to exploring results, offering a series of actions and shortcuts to control vertical and horizontal zooms for easy movement through analysis results. The tooltip window displays results corresponding to a given time (and position on the vertical axis when relevant). The editing mode enables you to modify analysis results directly on the graphical representation, by drawing curves, creating markers, copying and pasting results, etc. These operations can also be carried out textually via the results table. When analysis results are modified by the user, the data becomes

independent of the plug-in analysis and is saved in binary files associated with the document. This makes it possible not only to share the document with modifications, but also to undo and redo actions.

3.5. Exporting and sharing analyses

One of the major challenges of the project is to offer a tool that can be integrated into the existing uses and practices of users. To achieve this, Partiels offers interoperability with music creation software such as OpenMusic, Max, and Pure Data (or other digital audio workstations), as well as spreadsheet-type tools (or even simple text editors).

Partiels can be used to export analysis results in a variety of formats to suit different needs and contexts. It supports images in PNG or JPEG formats, and in addition to the usual size-related features, Partiels offers the option of exporting tracks independently or preserving group overlay. It also supports various text formats, including JSON, which offers the possibility of embedding analysis information (in order to restart the analysis afterwards, if necessary). It also offers support for the CSV format and its derivatives, which can be easily loaded into various software environments such as Pure Data, Max, Reaper[9], and others. This makes it possible, for example, to use this data to control synthesizers or, more simply, to segment audio files. More specific export formats, such as SDIF or CUE, are detailed in the user manual. It is also possible to export a single analysis, a group of analyses, or an entire document.

Partiels also offers a batch processing system with a graphical user interface that lets you apply a series of analyses to a set of audio files and automatically export the results in the desired format. In a development context, the command line interface can similarly be used to apply a set of analyses to an audio file and export the results. To facilitate the reproduction of analyses, it is possible to use an existing document as a template to apply its analyses to other audio files. This feature is available via the graphical interface or the command line interface.

Finally, Partiels offers the possibility of consolidating documents to facilitate the sharing of analyses and their results from one user to another, or from one computer to another. This action creates a folder associated with the Partiels document file, containing audio files and binary versions of the analysis results. So even if the computer or the user receiving the Partiels document does not have the plug-ins required for analysis, the results are still available.

4. ANALYSIS PLUG-INS AND EXTENSIONS

Partiels is based on Vamp plug-ins. It is compatible with all the analysis plug-ins already available[10] and distributed by numerous institutions such as Queen Mary University, the BBC, Pompeu Fabra University, and others. The project also includes several analysis plug-ins developed at IRCAM, which have required the creation of a Vamp extension.

4.1. IRCAM plug-ins

A great deal of work has gone into ensuring that the analyses are compatible with the main operating systems (macOS, Windows,

---

[9] https://www.reaper.fm/
[10] https://www.vamp-plugins.org/

< **301** >

and Linux), so that these tools are accessible to as many people as possible. This was particularly an important challenge for plug-ins based on machine learning technology[11], notably the TensorFlow[12] [19], a deep learning framework optimized for distributed training, and PyTorch [20].

Some of these plug-ins are based on technologies developed at IRCAM by the Sound Analysis and Synthesis team of the STMS laboratory:

- SuperVP: This plug-in offers numerous analyses based on the eponymous phase vocoder [1]. These include pitch estimators (for melodic and percussive instruments, and one based on the FCN neural network [21]), spectrogram analysis using FFT, LPC, Cepstrum, True Envelop, transient detection, marker generation using spectral difference detection, formant analysis, and voice cutoff frequency analysis.

- IrcamBeat: This plug-in allows you to estimate the tempo of an audio file and generate markers at each beat (which adapts to tempo fluctuations) [4].

- IrcamDescriptors: This plug-in offers four spectral descriptors (centroid, decrease, roll-off, and spread) and two perceptual descriptors (loudness and sharpness) [5].

- PM2: This plug-in offers harmonic and inharmonic partial detection with frequency, phase, and amplitude information, and a marker-based segmentation model for chord detection [2][3].

- VAX: This plug-in is based on neural network models for text-to-audio alignment. It displays a probability matrix of Latin characters present in the text and segments and realigns a given text [22], as shown in Fig. 4.

The other plug-ins are based on technologies external to IRCAM, such as:

- Crepe: This plug-in offers height estimation using a set of neural network models developed by Jong Wook Kim, Justin Salamon, Peter Li, and Juan Pablo Bello [7] and based on the TensorFlow framework [19].

- Whisper: This plug-in enables audio transcription using neural network models developed by OpenAI [6]. The integration is based on the Whisper.cpp C/C++ library developed by Georgi Gerganov[13] for optimized real-time inference.

- Basic Pitch: This plug-in enables multi-phonic pitch estimation using models developed by Spotify [23] and based on TensorFlow [19].

4.2. The IRCAM VAMP extension

When implementing the plug-ins, we came up against some limitations of the Vamp format. The Vamp plug-in SDK does not allow chaining of analyses, and only accepts scalar digital parameters and audio streams as input. This restriction prevents the development of complex approaches, such as aligning text to audio with the VAX engine, which requires text to be supplied upstream of the analysis. The IRCAM Vamp Extension (IVE)

---

[11]We managed to integrate these technologies in the plug-in as statically linked libraries to ensure the portability of the binaries and avoid the conflicts with third-party installations.

[12]https://www.tensorflow.org

[13]https://github.com/ggerganov/whisper.cpp

library, developed as part of this project, overcomes this constraint. It offers new interfaces to plug-ins, associated with those offered by the Vamp SDK, so that they can receive the results of other analyses in conjunction with the input parameters. In Partiels, this enables analysis chains to be created. For example, Whisper analysis is used to transcribe the text of an audio file, and the results are sent to VAX analysis to align the text with the audio. Updating an input plug-in parameter or modifying its results automatically triggers recalculation of the resulting analysis. Thus, changing the Whisper model or manually correcting the text restarts the analysis of the VAX plug-in and realigns the text to take account of the modifications.

A further restriction lies in the output formats of the analysis results as defined by the Vamp SDK. The result can be either a label, a scalar, or a vector of numerical values. In many cases, it is necessary to augment this output with additional numerical data to offer, for example, a confidence score for transients, or the amplitude and phase associated with the frequency of a partial. The IVE library offers the option of adding additional numerical data to the output data. In Partials, these values are displayed in addition to the main result data. They can also be used to modify representations by filtering results according to value thresholds. In this way, it is possible to hide certain partials whose amplitude is below a minimum threshold.

5. CONCLUSIONS

The Partiels project is now at an advanced stage of development. It offers a set of tools whose functionalities meet a large number of needs, contexts, and use cases, thanks to the analyses available via its numerous plug-ins. The software is already in use, both inside and outside IRCAM, in the fields of research and artistic creation, as well as in teaching, notably at Sorbonne University in musicology courses and the Sciences et Musicology curriculum.

However, there are still many aspects that can be completed and improved. One of the challenges is to continue to make this tool accessible to as many people as possible. In particular, we plan to offer video tutorials and translations of the interfaces (for the moment, the software is only available in English). On the other hand, we consider integrating certain IVE features directly into the Vamp SDK so that these improvements can be used in other Vamp hosts (Sonic Visualizer, Max, Pure Data, etc.) and so that IRCAM plug-ins are better supported outside Partiels. Generally speaking, IVE could be used as an experimental space before more stable integration within the Vamp SDK. As part of this experimental approach, we would like to integrate a graphical API into IVE to enable more specific representations for certain analyses, particularly when the results are multidimensional. In this regard, we are currently working on an adaptation of The Snail [24] analyses to generate a spectrogram from data where the color of a pixel depends on multiple data (and not just a single amplitude value). We also want to continue developing Max and Pure Data patches to offer alternative representations of the data and also to facilitate the sonification of this data. Finally, we are continually developing new analysis plug-ins based on new technologies that we are developing at IRCAM or on external models that use PyTorch in particular.

< 302 >

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] P. Depalle and G. Poirrot, "SVP: A modular system for analysis, processing and synthesis of sound signals," in *Proceedings of the International Computer Music Conference*, Montreal, Canada, 1991.

[2] P. Depalle, G. Garcia, and X. Rodet, "Tracking of partials for additive sound synthesis using hidden Markov models," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Minneapolis, USA, 1993.

[3] A. Röbel, "Estimation of partial parameters for non stationary sinusoids," in *Proceedings of the International Computer Music Conference*, New Orleans, USA, 2006.

[4] P. Peeters, "Template-based estimation of time-varying tempo," in *EURASIP Journal on Advances in Signal Processing, Volume 2007, Issue 1*, 2007.

[5] P. Peeters, "A large set of audio features for sound description (similarity and classification) in the cuidado project," Tech. Rep., IRCAM, Paris, France, 2004.

[6] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust speech recognition via large-scale weak supervision," in *https://arxiv.org/abs/2212.04356*, 2022.

[7] J. W. Kim, J. Salamon, P. Li, and J. P. Bello, "CREPE: A convolutional representation for pitch estimation," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Alberta, Canada, 2018.

[8] N. Bogaards and A. Röbel, "An interface for analysis-driven sound processing," in *in AES 119th Convention*, New York, USA, 2004.

[9] N. Bogaards, A. Röbel, and X. Rodet, "Sound analysis and processing with AudioSculpt 2," in *Proceedings of the International Computer Music Conference*, Miami, USA, 2004.

[10] M. Puckette, "Pure Data: another integrated computer music environment," in *Proceedings of the International Computer Music Conference*, Thessaloniki, Greece, 1997.

[11] J. Bresson, C. Agon, and G. Assayag, "OpenMusic – visual programming environment for music composition, analysis and research," in *Proceedings of the 19th ACM international conference on Multimedia*, Scottsdale, USA, 2011.

[12] P. Boersma and V. van Heuven, "Praat, a system for doing phonetics by computer," in *Glot International*, 2001, vol. 5, p. 341–345.

[13] P. Couprie, "iAnalyse : un logiciel d'aide à l'analyse musicale. journées d'informatique musicale," in *Actes des Journées d'Informatique Musicale*, Albi, France, 2008.

[14] C. Cannam, C. Landone, and M. Sandler, "Sonic Visualiser: An open source application for viewing, analysing, and annotating music audio files," in *Proceedings of the ACM Multimedia 2010 International Conference*, Firenze, Italy, 2010.

[15] C. Cannam, E. Benetos, M. E. P. Davies, S. Dixon, C. Landone, M. Levy, M. Mauch, K. Noland, and D. Stowell, "MIREX 2015 entry: Vamp plugins from the centre for digital music," in *Proceedings of the International Symposium on Music Information Retrieval*, Urbana-Champaign, USA, 2015.

[16] C. Cannam, C. Landone, M. Sandler, and J. P. Bello, "The Sonic Visualiser: A visualisation platform for semantic descriptors from musical signals," in *Proceedings of 7th International Conference on Music Information Retrieval*, Victoria, Canada, 2006.

[17] M. Wright and A. Freed, "Open SoundControl: A new protocol for communicating with sound synthesizers," in *Proceedings of the International Computer Music Conference*, Thessaloniki, Greece, 1997.

[18] C. Frauenberger, *The SuperCollider Book*, MIT Press, 2011.

[19] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: A system for large-scale machine learning," in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16)*, Savannah, USA, 2016.

[20] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "PyTorch: An imperative style, high-performance deep learning library," in *Conférence: Advances in Neural Information Processing Systems (NeurIPS 2019)*, Vancouver, Canada, 2019.

[21] L. Ardaillon and A. Röbel, "Fully-convolutional network for pitch estimation of speech signals," in *Insterspeech 2019*, Graz, Austria, 2019.

[22] G. Doras, Y. Teytaut, and A. Röbel, "A linear memory CTC-based algorithm for text-to-voice alignment of very long audio recordings," *Applied Sciences*, vol. 13, no. 3, 2023.

[23] R. M. Bittner, J. J. Bosch, D. Rubinstein, G. Meseguer-Brocal, and S. Ewert, "A lightweight instrument-agnostic model for polyphonic note transcription and multipitch estimation," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Singapore, 2022.

[24] T. Hélie and C. Picasso, "The Snail: A real-time software application to visualize sounds," in *Proceedings of the 20th International Conference on Digital Audio Effects (DAFx-17)*, Edinburgh, UK, 2017.

< **303** >