



HAL
open science

Audiomatching for the philological analysis of electroacoustic music

Nicola Orio, Laura Zattra

► **To cite this version:**

Nicola Orio, Laura Zattra. Audiomatching for the philological analysis of electroacoustic music. 2007
International Computer Music Conference, Aug 2007, Copenhagen, Denmark. hal-03139717

HAL Id: hal-03139717

<https://hal.science/hal-03139717>

Submitted on 12 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AUDIO MATCHING FOR THE PHILOLOGICAL ANALYSIS OF ELECTROACOUSTIC MUSIC

Nicola Orio

Laura Zattra

University of Padova
Department of Information
Engineering

University of Padova
Department of Visual Arts
and Music

ABSTRACT

The analysis of electroacoustic music may benefit from the design and development of tools for automatic audio matching, allowing musicologists to compare different versions of the same work and to draw the artistic path from the synthesis of fragments to the creation of a complete work. Music alignment has been explored as a tool in many areas, from score following to music information retrieval. This paper presents a tool for the analysis of electroacoustic music based on the comparison of two audio sources. The tool has been applied to a relevant case study: *Stria* by John Chowning.

1. INTRODUCTION

This paper investigates how automatic analysis and matching can provide the musicology with useful tools for the analysis and the philological study of electroacoustic music. It also stresses the urgency for the musicologist and the sound engineer to collaborate in order to find new analytical methods deriving from the combination of their competencies.

This research starts from two premises. The first one is the necessity, from a musicological point of view, to apply methods borrowed from the philology of music to the study of electroacoustic music. The comparison of different sources helps and reinforces the listening analysis based on visualization tools such as spectrograms and sonograms. But sources in electroacoustic music are very heterogeneous: as well as the tape or the CD, they also consist of notations, projects and schemes, digital scores (in the computer music field), live electronics instructions, comments for the performers, notes, software, various articles, reviews, and so on. However, although in different forms compared to traditional Western music, electroacoustic music stays anchored to the writing techniques and tools: texts, (analog and digital) scores, supports, and notation [2]. Musicological investigation should start from this idea and could set its activity within the research field named “philology of author” [28]. The philology of author, as proposed in [14], widely explores the revision activity by the composer, in the light of his/her interrelation with personal aesthetics and cultural and biographical

context¹.

The second premise emphasizes the problem of electroacoustic music archives: historically, the urge of creation and experimentation with electroacoustic (analog or digital) devices, has brought composers and their collaborators to ideally postpone, but often to forget, the organization and preservation of their compositional materials. It frequently happens, for example, that audio materials stored in many electroacoustic music studios do not contain the piece indicated on the label, or are illegible. Sometimes, one can find tapes with the same title but different audio contents, belonging to various steps of the realization process. It can also happen that the piece a composer considers as his/her final version slightly differs, for some reasons, to the published one.

The idea of this actual research comes from the musicologist’s need to find useful tools for the philological comparison of the audio content within the study of the electroacoustic music pieces. This necessity was brought to light during the study of the musical archive belonged to Teresa Rampazzi, pioneer of Italian electroacoustic music [27]. As it was this composer’s habit to make many copies and sometimes different versions of her pieces without writing any comments on the documentation (except for the title on the tape-box), it has become urgent to find a functional method for the verification of various audio materials, for their comparison and their investigation.

This paper describes the first steps of this research in finding a convenient method and application of an automatic tool for the analysis and comparison of different audio sources. Section 2 gives an overview of the different approaches to the matching of alternative versions of the same music work, while Section 3 describes the proposed system. However, this paper does not show the applications of the tool on Teresa Rampazzi’s repertory. It rather describes its setting up and verification applied to another piece: *Stria* (1977) by John Chowning. Section 4 describes the main characteristics of the piece and

¹ The need for embracing philology of authors comes not only from the fact that the youngest musical genre made with digital techniques benefits of all these types of documentation, but also because it is already facing a serious sustainability problem: while its history is indeed considered very recent, several technological generations have gone in the meantime [3]. Philology of author could help to restore the creative process.

the results obtained thanks to the philological analysis described in [29]. The results achieved through the use of the analysis tool are presented in Section 5.

The object of this research is the so-called fixed music. This is electroacoustic music which is fixed on a recording medium (e.g. analogue or digital supports, included the same computer used both as a storage support or as the medium by which the music could be calculated) and played back in performance over an amplified system. "Fixed" means generally that the work is finished when it leaves the studio, but here it also means a fixed intermediate level before the ultimate version.

2. RELATED WORK

The approaches to automatic music alignment can be classified according to the forms taken into account, audio recordings and symbolic scores, or to the aims of the alignment.

2.1. Aligning Audio and Symbolic Representations

Audio to score alignment is motivated by the central role played by the symbolic representation in Western music. The score representation is the main tool used by musicologists and music theorists in their research, giving a strong emphasis on the musical parameters that are easily and directly represented in symbolic notation, in particular pitch intervals and note durations. To this end, it has been noted that for Western art music there is a strong tendency to consider the symbolic representation as the ideal version of a musical work, to which performances are only approximations. On the other hand, a music score can be considered itself an approximate representation of a musical work, because it is impossible to represent all the different nuances of musical gestures with a compact symbolic representation [18].

Bridging the gap between the audio and symbolic representations is a research field that is related both to music information retrieval (MIR) and to computer music research communities. In the case of MIR, audio to score alignment is not a task by itself, but rather a preprocessing step for retrieval and recognition tasks. Music retrieval of symbolic score has been the subject of scientific research since more than a decade [11], because the application of retrieval techniques to scores can take advantage of a compact and direct representation of the music content. Aligning audio to scores allows a system to index and retrieve scores while providing recordings to the users as the results of a search task. Moreover, the alignment of different recordings with the same music score may be a valuable tool for the analysis of performers, for instance comparing alternative expressive timings [4]. Another application is score driven segmentation of the music signal, aimed at the extraction of isolate tones and chords that can be used for concatenative synthesis [23].

Approaches for audio to score alignment have to deal with the fact that, at least one of the two forms has to be

transformed in order to compute a match with the other. For example, the information in the score can be used to create a set of filterbanks, where each filter is centered on the expected harmonics of the signal, and the local match can be computed by measuring the energy that is output by the filterbank [24]. Alternatively, the concept of filterbank can be substituted by a modeling of the main statistical parameters of the expected harmonics [22]. The score can also be used to create an artificial performance, that is then matched against the real performance [26].

Once a mechanism for local matching is provided, the global alignment can be computed using well-known techniques for approximate matching and statistical modeling. The two most used techniques are Dynamic Time Warping (DTW) and Hidden Markov Models (HMM), both of them quite popular in the speech recognition research area [21]. An example of an application of DTW is presented in [13], while HMMs have been proposed in [19]. In both cases, alignment has been applied to an information retrieval task, and thus it could be carried out using the complete information on the performance and on the score. A more difficult task is the on-line alignment, which is presented in the following section.

2.2. Score Following

One of the main applications of automatic audio to score alignment is usually called score following. In this case a local alignment is computed between an ongoing performance, which is digitized and processed in real time, and a digital score stored in the system. The goal is to perform an automatic accompaniment able to follow the time deviations of the performance, and possibly to resynchronize in case of errors made by the musicians. Applications range from tools for instrumental practising to complete systems for public performances.

Early score following systems were based on dynamic programming approaches, such as [8], using MIDI format to represent both the score to be followed and the ongoing performance. It is interesting to notice that, already in 1993, an information retrieval approach has been applied to a score following task [25], where dynamic time warping were used on MIDI signals. While early approaches focused on the alignment of a monophonic solo instrument, and were based on external pitch trackers, more recent approaches directly deal with the audio signal, using statistical models [12], hidden Markov models [22, 5, 20], and more recently hierarchical hidden Markov models [6].

2.3. Aligning Two Alternative Audio Representations

Audio to score alignment is based on the availability of a digital representation of the score. While this assumption is obviously true for score following, there are situations in which it could be easier to create a database of models for the recognition using a collection of audio recordings. Moreover, the alignment of two alternative performances of the same musical work allows the musicologists to compare the style of different performers, and

can be exploited for resynthesizing performances with different expressive parameters. Audio to audio alignment may be the only option for genres that are not commonly represented by a standard notation, such as ethnic or electroacoustic music.

Audio to audio matching is usually based on a preprocessing of the recordings in order to extract relevant features that are able to generalize their main characteristics. A popular descriptor is the chroma-based representations, which has been proposed in [10]. The basic idea is that all the components of the spectrum are conflated in a single octave, obtaining a particular signature of a polyphonic signal. Alternatively, as presented in [15], audio recordings can be segmented in coherent parts with stable pitch components, and a set of bandpass filters are computed for each segment around the main peaks in the frequency domain.

Once a set of descriptors is computed from the two audio signals, the global matching can be carried out with the same techniques described in Section 2.1, in particular DTW and HMMs. For example, a variant of DTW has been proposed in [17] for off-line alignment, while a real time version of DTW has been presented in [9]. An approach to alignment based on HMMs is described in [15].

A particular application of audio to audio alignment regards electroacoustic music. In this case, the algorithms for the synthesis may not be available or difficult to reproduce without the systems that have been used by the composers. Moreover, the final recordings may be the results of post processing and audio editing, for which there is no available information. The aim of audio to audio matching for electroacoustic music is the analysis and comparison of alternative recordings, assembling, and mixing. In this case, audio to audio matching should be considered as a tool for the musicologist, who should be able to inspect local and global matching according to a number of parameters that could be changed during the analysis.

3. METHODOLOGY

3.1. A Tool for Comparing Recordings

From the discussion in Section 2 it is clear that there are a number of approaches that can be applied to audio matching. In particular, regarding the alignment problem, from the analysis of the literature it appears that there are two main approaches: Dynamic Time Warping (DTW) and Hidden Markov Models (HMMs). In both cases, the global alignment is computed from local distance using a dynamic programming approach. The main difference is that HMMs require that a model is built from one of the recordings, while DTW can be carried out directly from the signal parameters without giving a *preference* to one of the recordings. This is the main reason why DTW has been used in our research to compute the alignment. Another reason is that, for the musicologists who analyze the results of the automatic alignment, it is more intuitive to think about distances rather than marginal probabilities.

Particular attention has been paid to the visual representation of the audio matching. The aim of the proposed system is to provide the musicologist with a useful tool for analyzing and comparing electroacoustic music. Thus the first step is to compute and visualize the similarities of each couple of audio frames in the two recordings. Given the relevance of spectral information, the similarity function is based on the frequency representation of the signal. The user may choose the degree of granularity of the analysis, by selecting the length of the audio frames to be Fourier transformed, the hopsize between subsequent frames, and the final resolution in milliseconds.

An important choice regards the way the similarity between frames is computed. There are many approaches for the computation of a similarity matrix. Also in this case, our approach is to provide the users with tools, rather than selecting one single approach. Hence, the system allows the user to choose among a variety of distance/similarity functions. As expected from the analysis of the literature, the similarity function that gave the most interesting results, at least for the case study reported in Section 5, is the cosine of the angle between the vectors representing the amplitude of the Fourier transform, which has been used for the visualizations of the figures depicted in this paper.

As it is well known, DTW computes the best aligning path $p(m, n)$ across the local distance matrix d through a minimization, for instance according to equation

$$p(m, n) = \min \left\{ \begin{array}{l} p(m-1, n-1) + 1.5d(m, n) \\ p(m-1, n) + d(m, n) \\ p(m, n-1) + d(m, n) \end{array} \right\} \quad (1)$$

even if there have been proposed many different combinations of neighbor points to compute the minimization. The results presented in Section 4 have been computed using this equation, which is based on just three neighbors, located on a square.

We allow the user to choose the set of neighbor points used to compute the alignment, including points on a square, a triangle, and more complex figures, with selectable size. This way the user may choose the configuration that gives a more visually clear and musicologically correct alignment. Another important characteristic, that probably does not make sense for other music genres, is that the alignment can be computed also when one of the recording is reversed². In this case, the user chooses a starting point for the alignment and the system verifies all the possible directions in the time axis.

A Matlab application has been developed, and used to carry out a musicological analysis on a case study. A screen dump of the application is shown in Figure 1, reporting the alignment of two complete versions of *Stria*.

² This could be useful for the future analysis of the work by Teresa Rampazzi, who sometimes used to reverse tapes when recording her pieces.

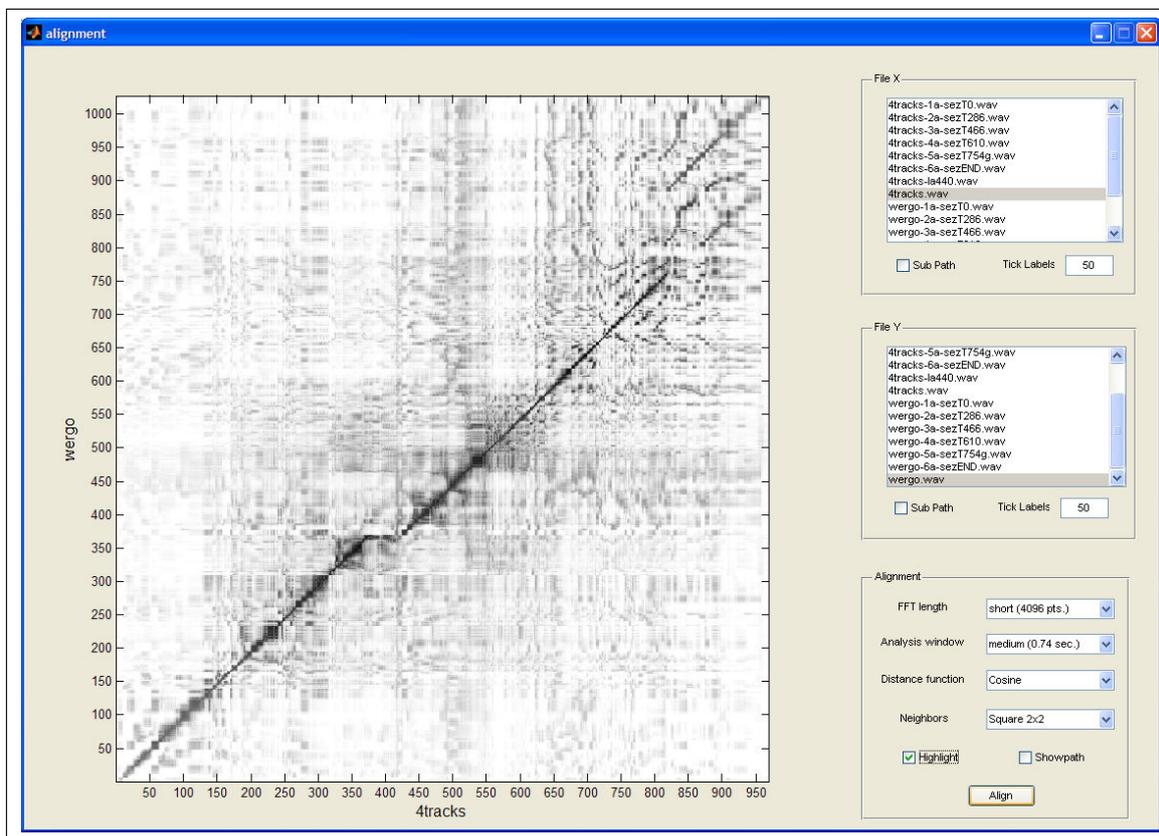


Figure 1. Screen dump of the Matlab application for the alignment of WERGO and CCRMA versions of *Stria*

4. CASE STUDY

The application of audio to audio alignment for electroacoustic music research has been tested on a historical piece of computer music, *Stria* by John Chowning. *Stria* exists in several versions and the genesis of the assembling of its sections is shown in a forthcoming article [29]. The results of the philological analysis are useful because they permit us an accurate reading and understanding of the plots resulting from the application of the audio matching tool.

It is also worth noting that the philological method, previously developed in [28], had followed different steps: information about the musical works, its history, main criteria and problems the analysis encounters during the philological instigation; complete *recensio* of extant sources; sources criticism (complete survey of the witnesses and description, list of all different sources, labeling with abbreviations derived from their content or origin, bibliographical or recording references or name of the archive); accurate description of the sources; systematic collation (comparison); restitution of the text (central and conjectural work of the philologist). This methodology is shown in [29]. The aim of the philology of author is to make a critical investigation of the musical piece and its tradition, trying to restore its history, innovations and changes during its transmission and reception. Results depend on the accuracy of these steps.

Through the course of our past research, it had been

evident that *Stria* was developed in different phases and had different durations since its first realization in 1977 and his publication in the Digital Music Digital Series by WERGO³ in 1988.

Stria was realized in sections, synthesized, converted and recorded from the computer PDP-10 onto 4-channels analogue tapes (one for each section). Thanks to the philological investigation on the history of the synthesis and the various stages of assembling of its sections, three main versions of the piece did emerge: the IRCAM version (this is conjectural because the tape is missing) played during the first performance⁴; the CCRMA version, made the day following the premiere; the WERGO recording, made in 1988.

The discrepancies concern the overall duration but depend on the use or removal of certain sections or their shortening. Following our previous analysis, we know that these versions use identical materials: the same tapes were assembled, since at CCRMA they did not re-compute the sections because of the massive amount of the computational effort involved⁵. Johannes Goebel, composer, producer, and technician who worked in 1988 for WERGO, transferred the data from the original sound files from the mass storage disk and kept them on digital tapes.

³ John Chowning, Turenas; *Stria*, Phoné; Sabelithe. Digital Music Digital, Music with computers. Mainz: Schott – Wergo Music Media GmbH. WERGO WER 2012-50.

⁴ October 13, 1977 at Centre Pompidou.

⁵ John Chowning personal communication.

Version	Date	Duration	Comments
CCRMA	1977-1978	15'46"	Sections T0 (166") – T286 (252" version with discontinuity) – T466 (144") – T610 (144") – T754G (short version, 111") – END (145")
WERGO	1988	16'57"	T286 third truncation (206"): Chowning truncated the discontinuity T754G (long version, 234") Compromise in dynamic range

Table 1. Main characteristics of the assembling (adapted from Table 4 in [29])

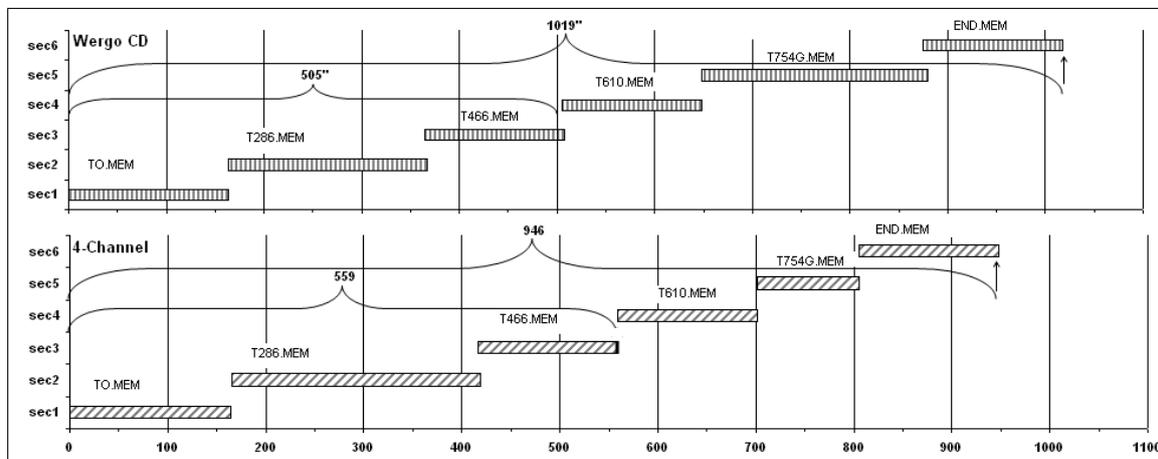


Figure 2. A graphical representation of the assembling of *Stria* by John Chowning (adapted from Figure 16 in [29])

The listening analysis and its comparison with the graphical representation (amplitude/time representation and sonogram) concentrated on CCRMA and WERGO versions. The analysis located the beginnings of each section and made us know how each section sounds and how it is mixed with its following. Those results were compared with the corresponding digital sources (sources scanned from paper printout of the digital data used for the synthesis and the calculation of the structure of the piece, John Chowning Papers, Stanford University Archives) and the explications by the composer himself.

The analysis established that: TO.MEM, T466.MEM, END.MEM were equal sections in both versions; while T286.MEM and T754G.MEM had different durations⁶. Figure 2 shows the assembling of the two versions. One can clearly notice that CCRMA and WERGO second and fifth sections have different durations.

The reason of the truncated part of T286 in the WERGO version was caused by the decision to cut an unintended “discontinuity” in the D/A conversion of the data occurring at 6'29". This problem caused a sudden change of timbre which conflicted with the constant flowing of the FM spectra (this “faulty” CCRMA version is the version Chowning nowadays uses to play during live performances with spatialization).

The reason of the extended fifth T754G section in the WERGO version remains obscure. Probably the originally converted T754G, had been considered too long and cut by the composer for the premiere, but the truncation for the WERGO was ignored. Table 1 explains these dis-

⁶ MEM files were records of data used for the generation of each section of the piece. The number after T indicates the time of initialization, even if the philological analysis outlines a series of alterations and inaccuracies made during the compositional process.

crepancies.

Since we want to validate these results through audio matching, CCRMA and WERGO audio files needed some adjustments in order to fulfill the Matlab application requirements, which has been developed to match monophonic files. CCRMA version is taken from a recent digitalization of the 4-channels analogue tape⁷. We have mixed the 4 tracks down to mono. Moreover, we have cut the first seconds of the track which contains four sounds (Test Tones) for testing right and left, front and back loudspeakers for performance in concert, whose presence is irrelevant for our purposes. The WERGO version is taken from the CD published in 1988. The track was stereophonic and has been mixed down to mono. Later, we have also divided the overall tracks in several sections, following the assembling organization of the piece. Since the piece is characterized by a constant flowing of the spectra and the assembling is made by overlapping the sections, it is sometimes hard to cut the sections when they end and/or start. Therefore for each sub-section we have always kept the last seconds of the precedent section and the first seconds of the following.

5. EVALUATION

The first audio matching was made using the complete CCRMA (indicated as *4tracks* in the figure) and WERGO versions, and are reported in Figure 1. The figure clearly shows the common sections. At 375" or so, something happens. This corresponds to the moment the discontinu-

⁷ The digitalization was made on November 2002 by Alvis Vidolin. The CD-Rom was kindly provided by the Italian composer Roberto Doati.

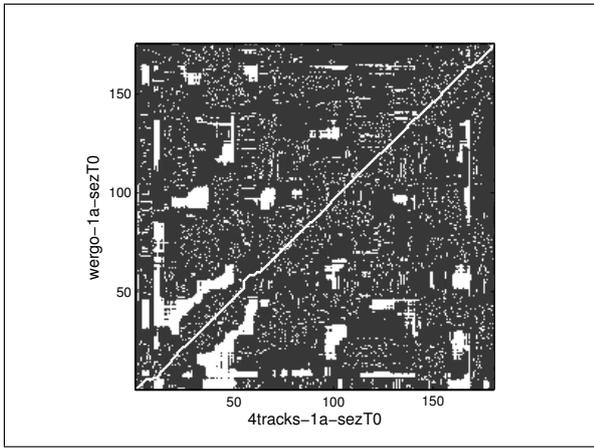


Figure 3. Alignment of section T0

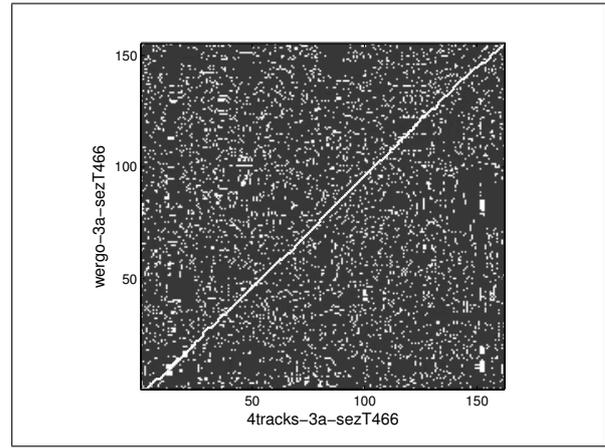


Figure 5. Alignment of section T466

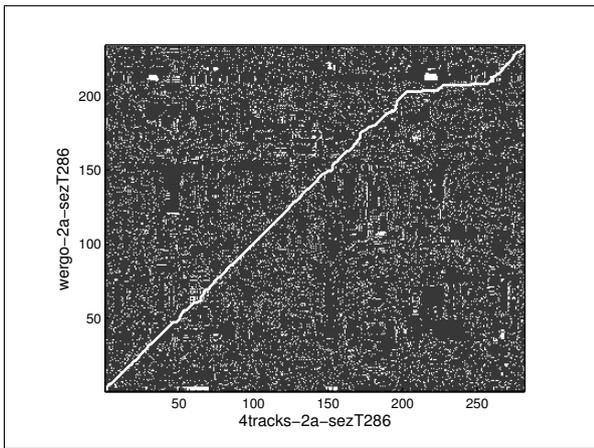


Figure 4. Alignment of section T286

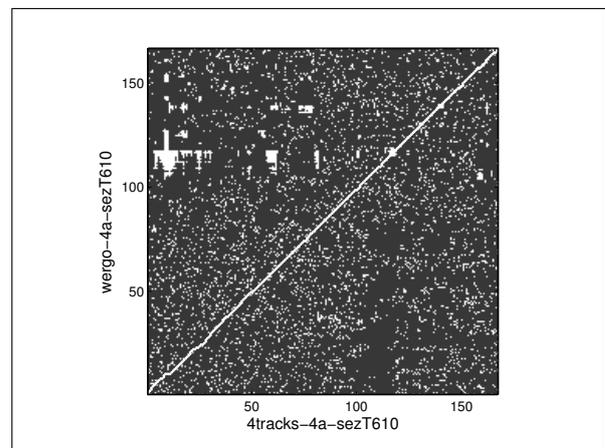


Figure 6. Alignment of section T610

ity occurs (T286) in the 4-tracks-CCRMA version (6'29", that is 389"). The path changes from 350" to 420" because T286 in the CCRMA version is longer, whereas T286 in the WERGO version had been truncated right before the discontinuity. After that, the two audio files match again.

Another controversial point happens at 810-20". This is the moment where the two T754G differ in duration: WERGO CD uses the long section, whereas CCRMA version was made with the short T754G section: the END section immediately follows to it. From here to the end the spectral space shows different possible alignment. This is due to the particular spectral nature of the last two sections (T754 and END). It is worth to mention that T754 was the original end, but as Chowning recalls, before leaving the first performance, he added the END section. These two sections have actually similar spectral contents and both end on C \sharp 7⁸. An accurate analysis of the piece, section vs. section, emphasizes many other details. It is interesting to notice that not only the alignment confirms the musicological analysis, but it also shows from time to time new details which escaped the analyst: these new discrep-

⁸ END was also one of the first sections the composer generated, and it did not have that name. He finally decided that it was a perfect ending and he renamed it END.MEM.

ancies could be caused by the conversion from 4-channels and stereo channels to mono, or by practical adjustments in the dynamic range that were made for the CD for domestic listening (some soft parts of the WERGO CD are in fact much louder), or by spectral differences in the two versions that could bring to light additional philological problems.

T0 alignment confirms the fact that these two versions did use the same original tapes (Figure 3), even if at 55" the slight differences in the two recording induce the algorithm to choose an alternative path during a stable spectrum. T286 sections, shown in Figure 4, have the same FM spectra up to 200", then the discontinuity shown in the global alignment is even more clear. The similarity starts again at 260" when the following section of the piece, that is T466, begins. T466 and T610 sections do use the same materials; their files correspond and show no particular problems. For completeness they are shown in Figure 5 and in Figure 6 respectively.

CCRMA version employs the beginning of the T754G section. T754G lasts 234 (3'54"), as marked in both WERGO and in the digital score of the piece [29]. The short T754G used in 4-channels version is a version that was cut down at a certain point of the realization. Figure 7 does not show

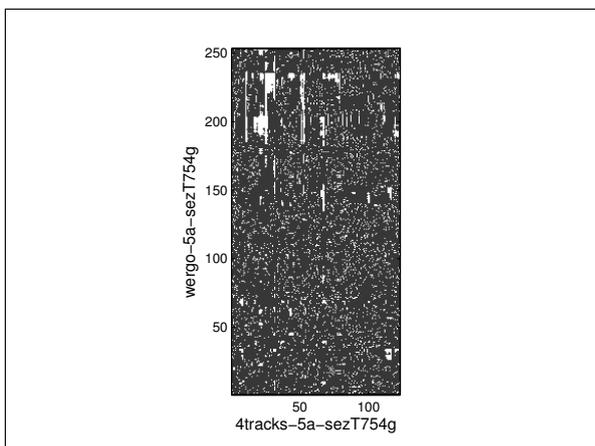


Figure 7. Alignment of section T754G

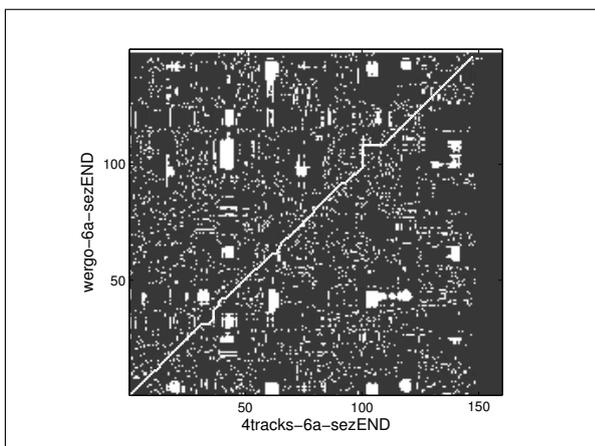


Figure 8. Alignment of section END

the path in order to better highlight the spectral image of the two sections. END sections (Figure 8) are a bit problematic. They have the same duration and have the same spectral content, but the angle bisector shows some problems at 30", and at the end of the section. This is very interesting since, theoretically, they would use the same tapes. Our philological analysis did outline - and was confirmed by the documentation's and Chowning's testimony - that the sections were never re-calculated. Therefore these new discrepancies force us to deepen our future investigation.

6. CONCLUSIONS

Audio matching is useful for the analysis of electroacoustic music because it deals with a music characterized by complex spectra. *Stria* is a significant example of complex music (constant flowing of FM spectra) which is particularly difficult to analyze by pure listening⁹ and requires a great experience and specific tools for the detailed FFT analysis [7]. Even if, listening analysis managed to draw the macro analysis of different sections [29], it was im-

⁹This did not prevent Toby Mountain to make a beautiful analysis in 1980 [16].

possible to detect micro inner changes - if they were - between the paired sections of CCRMA and WERGO version. Philological analysis of the digital score and other sources told us that the tapes used for the two CCRMA and WERGO versions were the same but little discrepancies in the audio to audio alignment make us aspire to further research. Those discrepancies could hide other problems in the original files used for the assembling: noise or other sounds introduced by the transfer from the mass storage disk (Chowning recalls that maybe was Librascope) on digital tapes in 1988 when the CD was made, or other sort of problems, which could add interesting philological consequences. The re-synthesis of the piece made by Kevin Dahan [7] and Olivier Baudouin [1] could help answering these questions.

Further research should improve the algorithm with new tools for the musicological analysis. These could be the possibility either to inquire the representation from the spectral point of view, e.g. with pop-up windows indicating FFT analysis where the cursor is pointed, or to make inquiries within the representation with details regarding the timing.

This tool is thought for the philological analysis of electroacoustic music. Its application could ranger from the philological study of different versions of electroacoustic music pieces, to the study of the genetic process of one piece, particularly when there are different audio 'sketches' which lead to the final version, and, in a more general level, to the study of the musical style in any musical genres. The object of the actual research is the so-called fixed music, but could also be improved for live music.

Acknowledgments

This research is partially supported by a Research Grant at the Department of Visual Arts and Music, University of Padova (L.Zattra) and by the SAPIR project, as part of the Information Society Technology (IST) Program of the European Commission (Contract IST-045128) (N.Orio).

7. REFERENCES

- [1] O. Baudouin. A Reconstruction of *Stria*. *Computer Music Journal*, 31(3), Fall 2007. forthcoming.
- [2] A.I. De Benedictis. Scrittura e supporti nel Novecento: alcune riflessioni e un esempio (*Ausstrahlung* di Bruno Maderna). pages 237–291. ETS, Pisa, 2004.
- [3] N. Bernardini and A. Vidolin. Sustainable live electro-acoustic music. In *Proceedings of Sound and Music Computing*, 2005.
- [4] S. Canazza, G. De Poli, C. Drioli, A. Rodà, and A. Vidolin. Modeling and control of expressiveness in music performance. *Proceedings of IEEE*, 92(4):686–701, 2004.

- [5] P. Cano, A. Loscos, and J. Bonada. Score-performance matching using hmms. In *Proceedings of the International Computer Music Conference*, pages 441–444, 1999.
- [6] A. Cont. Realtime audio to score alignment for polyphonic music instruments using sparse non-negative constraints and hierarchical hmms. In *IEEE International Conference in Acoustics and Speech Signal Processing*, pages V245–V248, 2006.
- [7] K. Dahan. Surface tensions: Dynamics of *Stria*. *Computer Music Journal*, 31(3), Fall 2007. forthcoming.
- [8] R.B. Dannenberg and H. Mukaino. New techniques for enhanced quality of computer accompaniment. In *Proceedings of the International Computer Music Conference*, pages 243–249, 1988.
- [9] S. Dixon and G. Widmer. MATCH: a music alignment tool chest. In *Proceedings of the International Conference of Music Information Retrieval*, pages 492–497, 2005.
- [10] T. Fujishima. Realtime chord recognition of musical sound: a system using common lisp music. In *Proceedings of the International Computer Music Conference*, pages 464–467, 1999.
- [11] A. Ghias, J. Logan, D. Chamberlin, and B.C. Smith. Query by humming: Musical information retrieval in an audio database. In *Proceedings of the ACM Conference on Digital Libraries*, pages 231–236, 1995.
- [12] L. Grubb and R.B. Dannenberg. A stochastic method of tracking a vocal performer. In *Proceedings of the International Computer Music Conference*, pages 301–308, 1997.
- [13] N. Hu, R.B. Dannenberg, and G. Tzanetakis. Polyphonic audio matching and alignment for music retrieval. In *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 185–188, 2003.
- [14] D. Isella. *Le carte mescolate. Esperienze di filologia d'autore*. Liviana, Padova, 1987.
- [15] R. Miotto and N. Orio. Recognition of music performances through audio matching. In *Proceedings of the Italian Research Conference on Digital Libraries*, 2007.
- [16] T. Mountain. A notational representation of John Chowning's *Stria*. handwritten graphical analysis. John Chowning Papers, Stanford University Archives. Publication forthcoming as a scanned image on the DVD-Rom, portion of Computer Music Journal Sound and Video Anthology Vol. 31, 2007, the DVD accompanying Computer Music Journal 31(4), 1981.
- [17] M. Müller, F. Kurth, and M. Clausen. Audio matching via chroma-based statistical features. In *Proceedings of the International Conference of Music Information Retrieval*, pages 288–295, 2005.
- [18] J.-J. Nattiez. *Musicologie générale et sémiologie*. Christian Bourgeois éditeur, Paris (FR), 1987.
- [19] N. Orio. Alignment of performances with scores aimed at content-based music access and retrieval. In *Proceedings of European Conference on Digital Libraries*, pages 479–492, 2002.
- [20] N. Orio and F. Déchelle. Score following using spectral analysis and hidden markov models. In *Proceedings of the International Computer Music Conference*, pages 125–129, 2001.
- [21] L.R. Rabiner and B.-H. Juang. *Fundamentals of Speech Recognition*. Prentice-Hall, Englewood Cliffs, NJ, 1993.
- [22] C. Raphael. Automatic segmentation of acoustic musical signals using hidden markov models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(4):360–370, 1999.
- [23] D. Schwarz. New developments in data-driven concatenative sound synthesis. In *Proceedings of the International Computer Music Conference*, pages 443–446, 2003.
- [24] F. Soulez, X. Rodet, and D. Schwarz. Improving polyphonic and poly-instrumental music to score alignment. In *Proceedings of the International Conference of Music Information Retrieval*, pages 143–148, 2003.
- [25] D.R. Stammen and B. Pennycook. Real-time recognition of melodic fragments using the dynamic time-warp algorithm. In *Proceedings of the International Computer Music Conference*, pages 232–235, 1993.
- [26] R.J. Turetsky and D.P.W. Ellis. Ground-truth transcriptions of real music from force-aligned midi syntheses. In *Proceedings of the International Conference of Music Information Retrieval*, pages 135–141, 2003.
- [27] L. Zattra. Teresa Rampazzi. Pioneer of Italian Electronic Music. In *Proceedings of the Colloquium on Musical Informatics*, pages 11–16, 2003.
- [28] L. Zattra. The critical editing of computer music. In *Electroacoustic Music Studies Conference (EMS06) – Theme : Terminology and Translation*, Beijing (CN), October 2006. http://www.ems-network.org/article.php3?id_article=236.
- [29] L. Zattra. The Assembling of *Stria*: A Philological Investigation. *Computer Music Journal*, 31(3), Fall 2007. forthcoming.