Team 1

Audio, Acoustical and Optical waves (AAO)

Head G. Richard (P)

Permanent staff R. Badeau (MC), B. David (MC), C. Févotte (CR2-CNRS, from 11/07), R. Frey (P, 40%), Y. Grenier (P), S. Maeda (DR CNRS), A. Maruani (P), D. Matignon (MC,-09/07), N. Moreau (P), S. Essid (IE, from 10/06), J. Prado (MC, on leave 06/07-09/08), I. Zaquine (MC, 80 %),

PhD students S. Essid (10/02-12/05), R. Badeau (10/01-04/05), D. Bitault (10/02-10/05), A. Aissa El Bey (10/04-06/07), S. Fontana (10/04-07/08), M. Guillaume (10/03-11/06), N. Bertin (10/05-), M. Betser (10/04-06/08), C. Clavel (11/03-03/07), V. Emiya (10/04-10/08), J-L. Durrieu (01/07-), O. Gillet (12/03-06/07), P. Leveau (11/04-11/07), M. Ramona (10/06-), M. Alonso (10/02-11/06), C. Baras (10/02-06/06), A. Moreau (10/03-09/06), Q. He (11/05-10/08), J-L. Smirr (01/07-), E. Ravelli (10/05-10/08), C. Joder (11/07-), L. Oudre (10/07-), F. Vallet (11/07-), S. Gulluni (02/08-), R. Hennequin (10/08-), M. Maazaoui (01/09-),

PostDocs, engineers and sabbaticals S. Essid (Postdoc 9 months), C. Hory (Postdoc 16 months), C. Févotte (Postdoc 8 months), M. Christensen (Aalborg Univ. (DK) Sabb. 1 month), A. Ozerov (Postdoc 18 months), M. Lagrange (Postdoc,10/2008-), T. Fillon (Postdoc,10/2008-), B. Mathieu (Engineer,10/2008-), Y. Menesguen (PostDoc 6 months).

External collaborators L. Daudet (Univ. Paris VI), O. Derrien (Univ. of Toulon), E. Vincent (IRISA), L. Devillers (LIMSI-CNRS), T. Sikora (Technical Univ. of Berlin) . . .

Permanent staff [*Institut ; CNRS*] ; post-docs	[7.5 ; 1.3] ; 1.5
PhDs	10
Defended PhDs	18
Defended HDR	2
Journal papers [published, in press]	[53 ; 5]
Chapters and books [published, in press]	[3 ; 7]
Conference papers	137
Patents and software	2
Contractual income 2005–2009 (june) [Private ; Public ; European] (k€)	[560; 755 ; 356]

The AAO (**A**udio, **A**coustical and **O**ptical waves) research group gathers digital and optical signal processing activities with a strong reference to the physical properties of the acoustical and optical phenomena. The group is structured in two research projects:

- Audio Signal Processing (*AudioSig* project),
- Optical Signal Processing (*TOS* project)

1.1 Audio Signal Processing (*AudioSig* project)

1.1.1 Objectives

The aim of this project is to develop digital audio signal processing methods in order to propose innovative solutions to the main problems linked to audio (speech, music,. . .) in multimedia applications. Our interests encompass the complete processing chain from sound capture and transmission to sound restitution. Work is both conducted on a methodological level to develop new sound representations and models especially for musical signals (Adaptive methods for high resolution sinusoidal components tracking, sparse representations, Non-Negative Matrix factorization, hierarchical models,. . .) and on their application to practical problems (watermarking, compression, EEG signal processing, automatic indexing). Audio indexing and retrieval currently is the central research theme of this project and includes topics such as broadcast streams segmentation into broad classes of audio events (speech/music/silence/singing,. . .), musical signals automatic analysis, decomposition and understanding (polyphonic audio source separation, rhythm extraction, multiple fundamental frequencies estimation, main melody extraction,. . .). A new transverse orientation has also gained more interest with the arrival in november 2007 of a new CNRS permanent researcher on the specific theme of statistical methods for audio signal processing.

On a different level, the group has initiated the development of a multimedia indexing and mining platform (called PLATO) which now involves several other groups. This internal platform, targeted to researchers, aims at being an intelligent media library, at centralizing research software, processing tools and computation resources and at providing demonstrative and communication tools.

The project is also maintaining tight links and collaborations with both academics (Queen Mary university of London, Dublin City University, Technical University of Berlin, University Paris 6 (LAM), IRCAM, INRIA-IRISA, LABRI-CNRS,. . .) and industry (Thalès, FT R&D, RTL, INA, Audionamix,. . .).

1.1.2 Results

Audio and multimedia scenes analysis and indexing

Researchers R. Badeau, B. David, S. Essid, C. Févotte, Y. Grenier, J. Prado, G. Richard;

Highlights :

Collaborations: With industry (FT R&D, Thales, RTL, INA) and academics (TU Berlin, Queen Mary University, LAM-Paris 6, IRISA, IRCAM, LABRI,...)
Projects: Network of Excellence IST-Kspace (*Knowledge Space of Semantic Inference for Automatic Annotation and Retrieval of Multimedia Content*), ACI Musicdiscover (*Indexing*

and search in audio databases), ANR-Desam (*Decompositions in sound elements and musical applications*), IVMN-infom@gic, ANR Sarah (*StAndardisation of High-Definition Remastering*, OSEO-QUAERO (*towards multimedia and multilingual search engines for professional and general public applications*);
Prize: PhD prize "ParisTech 2006" (R. Badeau)

This activity is following several research axes. The first direction, which is on a rather methodological level, aims at developing generic signal models and representations with a specific focus on audio signals. Several very interesting results were obtained for the estimation and tracking of sinusoidal components of an audio signal (new estimators for amplitude and frequency modulated components in noise [17], efficient algorithms for the adaptive estimation and tracking of the signal subspace components [9][12]). An increased effort was also dedicated to sparse signal representations, such as based on Matching Pursuit or Non-negative Matrix factorisation (NMF)[27], that allow to decompose a signal using a limited number of atoms or basis functions. The applicability of these methods to generic problems such as scalable audio signal compression [44], audio source separation or music signal indexing was demonstrated by introducing specific constraints deduced from the audio signal properties (use of instrument specific atoms for music instrument recognition [40], use of harmonicity or temporal constraints for music transcription[164], use of source production or timbre models for source separation [102],...). This methodological effort explores both deterministic and statistical approaches.

The second direction concerns the different facets of audio indexing and audio source separation which are two intricate problems. Indeed, efficient source separation eases the transcription of the resulting sources and efficient audio indexing facilitates the source separation. In music signal transcription, the group is directly interested in the four main problems which are *multiple fundamental frequencies estimation* (e.g. detection of simultaneous notes in a polyphonic musical recording [105],[164]), *rhythmical information tracking* (tempo and beat estimation [6, 5], *harmonic information estimation* (recognition of the chords sequence) and *timbre recognition* (musical instrument recognition in polyphonic audio [25]). Source separation approaches were developed for specific music transcription tasks such as drum track transcription and resynthesis [31]) and main melody estimation (by use of a NMF-based source-filter model for separating the singing voice from the musical accompaniment [103]) but also for specific audio rendering tasks such as stereo signal remastering [43].

The third research direction is dedicated to the audio streams segmentation into broad classes of audio events with application to *broadcast multimedia streams* (speech/music segmentation [157], speech emotion recognition [46],[21] or TV show structuring) and *musical streams* (musical instrument recognition [26],[38], multimodal audio/video semantic alignment [29]). Our efforts in this field is now evolving towards the automatic classification- both supervised and unsupervised- of multi-modal (or multi-stream) data sequences, typically audiovisual streams. Our emphasis is targeted to the incorporation of prior knowledge on the nature and structure of the streams processed, typically temporal dependencies and/or inter-stream correlations/dependencies, both at the signal level and the semantic level, possibly using ancillary information attached to the content (available meta-data, tags, notices, etc.) and/or user interaction (relevance feedback). At the methodological level, a special interest has been directed to kernel-based methods (Support Vector Machines, sequence kernels, probabilistic distances, kernel change detection, kernel LDA,. . .) and more recently to hybrid kernel and Bayesian network based methods.

Whenever possible, the results obtained are submitted to national or international evaluation campaigns. In particular in 2008, the group has participated to the national *ESTER 2 campaign* (Audio stream segmentation : best algorithm for music/non music detection and 2nd best for speech/non speech detection), the *Sissec campaign* (best results in two audio source separation subtasks) and *MIREX* (best algorithm for main melody estimation in 2008).

Sound capture and rendering

Researchers B. David, Y. Grenier, J. Prado, G. Richard;

Highlights Joint PhD with University of Parme, Italy; contract with France Télécom on audio source separation in the automotive domain, CapDigital-ROMEO (*a project within "pôle de compétitivité CapDigital, lead by Aldebaran Robotics and aiming at creating a humanoid robot*)

The objective of this theme is to improve sound field analysis and synthesis capabilities by developing specific digital signal processing methods. In binaural reproduction, a new approach was introduced to rapidly acquire new Head Related Transfer Functions (HRTF) and to personalize the rendering system to a new listener [113]. Such a binaural reproduction system, where the acoustics of a room are simulated as perceived by the listener through his HRTF, was developed. Formal perception tests were also conducted in collaboration with the university of Parme to validate the different sound rendering methods proposed [87].

In sound capture, recent work permitted to propose a novel technic for automatic sound field analysis from a network of sensors (microphones) [124]. This approach refers to the classical multi-microphones beamforming and parametric spectral estimation principles. The sound field component in each direction is obtained from the maximization of the spatial resolution around the targeted direction. This filtering is directly expressed under the form of spheroidal functions. Current work tackles the difficult problem of humanoid robot audition which needs, using a limited number of sensors, to be robust to movements of the robot and to highly variable environments.

Concurrently, a novel approach for blind audio source separation from a network of sensors was introduced for the underdetermined case (e.g. less sources than sensors). This method combines a wavelet-based time frequency analysis with an automatic classification of the data vectors that represent the positions of each source [4]. We produced several variants of this approach, one of them being based on an empirical modal decomposition [2]. We have shown that our blind separation techniques could be embedded in a general framework characterized by the use of second order statistical properties of the signals [3]. Since our goal was to apply these techniques in the car environment, we had to take into account the properties of the acoustic channels between the position of each source and the microphones (each channel acts as a filter or a convolution between the source signal and the impulse response of the channel); for this reason, another variant of the separation technique, which takes into account the convolutions, was elaborated in the time-frequency domain[1].

Sound sources watermarking and compression

Researchers N. Moreau, G. Richard

Highlights : Media Puppet project, academic collaborations (Univ. of Toulon, INPG Grenoble, Univ. of Paris 6/LAM)

Originally, the focus in audio watermarking was on the technology performances improvement (in terms of bit rates/ratio of binary errors) by introducing new methods exploiting the fact that a watermarking system can be viewed as a communication channel with adjacent information [14]. Recently, the objective was refocused on robustness issues to take into account typical use cases (such as those provided by Mediametrie). In particular, specific effort was dedicated to allow the detection of a hidden signal for degraded recordings (low quality microphones) or degraded communications (due to reverberation in a set-up where the loudspeakers and microphone are separated by at least 1m50). This appears to be a difficult problem that can only be partially solved by adaptive equalisation technics.

In audio compression, the work was mostly dedicated to low bit rate audio coding in the transform domain. On the one hand, specific effort was put to develop optimized quantization schemes for the MPEG Advanced Audio Coder (AAC) using a statistical subband model [22]. This approach was later extended to stereo signals for the MS-stereo mode of the AAC coder. In particular, the quantization error model introduced permits a global approach for coding both Middle and Side channels in the same process leading to improved efficiency without increase of complexity [23]. On the other hand, investigations were conducted to develop highly scalable transform coders which can seamlessly operate from very low bit rate up to transparency. To that aim, sparse overcomplete representations are used to decompose the audio signals over a redundant union of bases (such as Modified Discrete Cosine Transform bases at different scales) [44]. It was also shown that the high flexibility of the signal representations used in this coder allows to tackle various audio indexing tasks (such as beat tracking or musical genre recognition) directly in the transform domain [45].

Active noise control and biomedical signals analysis

Researchers J. Prado, Y. Grenier;

Highlights : External collaboration, ACI Abrupt (*Active Noise control of perceived background noise in call centers*)

In the framework of the ACI ABRUPT project, the activity focused on the development of appropriate methods for active noise control of background noise in call centers. For this purpose, a slightly modified GMDFa(Generalised Multi-Delay Filter) algorithm was used where the signal reconstruction by overlap and add was suppressed. Although this modification leads to slightly lower performances, it permits to obtain a lower complexity algorithm with still better noise suppression capabilities than time-domain approaches (such as FXLMS for example) especially in terms of signal processed bandwidth.

The other research direction is dedicated to the analysis of biomedical signals and especially electroencephalogram (EEG) signals recorded on asleep subjects using a single pair of sensors. Our approach to this problem has two technological breakthroughs since it aimed at an automated analysis (and not only visual) and uses a single channel EEG. The efficiency and robustness of the method developed have been measured and experimentally validated [175],[16]. The first goal of this method is to reduce the overall complexity (both in processing time and operation) of the standard approaches in order to obtain a hypnogram according to the rules of Rechtschaffen and Kales (R&K 1968) and that are adapted to the new rules of the American Academy of Sleep Medicine (AASM 2007). A hypnogram is a graphical representation of the sleep stages, from light sleep to deep sleep. Hence the method is able to control the drowsiness in real-time which has numerous industrial applications such as risky site monitoring or transport security (preliminary results are reported in [77]). Another direction of research targets the so call ”smart waking up” concept whose principle is to awaken a subject when the phase of sleep is the most favorable (light sleep or dream (REM stage)) to reduce the inertia of sleep. The sleep inertia is a transitional state of disorientation and confusion on awakening and may causes the degradation of mental performance. It was, in particular, shown that it is possible to optimize sleep to get the benefits (the recovery) without the disadvantages (torpor, sleep inertia).

Speech production

Researchers S. Maeda;

Highlights :
Collaborations: Collaboration With Department of Human Information Processing in ATR, Kyoto Japan and Phonetics and Phonology Laboratory (PPL), CNRS-University Paris 3.

Projects:IST-ASPI (Audiovisual to Articulatory Inversion), ANR-ARTIS (Articulatory inversion from audio-visual speech for augmented speech presentation), Experimental and Clinical phonetics with multi-instrumentations

In the context of the European project ASPI, we have investigated the acoustics characteristics of fricative sounds in various languages, which can be exploited in the acoustics-to-articulatory inversion. The combination of the high resolution MRI data recorded at ATR for the 3D vocal-tract shapes during the production of the fricatives and acoustic simulation have revealed that 1) distinctively different two classes of vocal tract configurations are used by French speakers to produce the same fricative consonant [160]; 2) a smooth change in the vocal-tract shape does not always produce a smooth spectral shape variation of the fricatives. Rather, in some regions the change produces a little spectral change whereas in other regions it causes an important spectral shape change. Interestingly the MRI observed vocal-tract shapes during fricatives tend to disperse in the stable regions, providing the evidence that the acoustic property of the vocal tract contributes to the specificity of the fricative sounds used in languages [142]; 3) we have developed relatively simple models of fricatives that can produce highly intelligible and naturally sounding fricatives in speech synthesis experiment [200].

In the follow up project, ARTIS, we are improving the acoustic modeling of fricatives and other consonants in order to fully exploit the advance in the MR imaging technique to measure detailed vocal-tract shapes. We expect that such modeling will allow us to gain the comprehensive understanding on the mapping between the vocal-tract shapes and the acoustic patterns of speech. The collaboration with Kiyoshi Honda (ATR) resulted in the invention of two non-invasive instruments: an external lighting and sensing PhotoGlottoGraph (ePGG) and a pneumotachograph with a disposal mask. The former is used to observe the activities of the larynx, abduction/adduction of the vocal folds during consonants and their oscillation during voicing. The latter one is used to measure the airflow passing through the vocal tract. These instruments will be used to evaluate the speech ability of patients in medical environments as well as in phonetic experiments [167]. Patent application for each of these two inventions is in progress with help from the CNRS.

1.2 Optical signal processing

Researchers R. Frey , A. Maruani , I. Zaquine ;

Highlights Institut TELECOM funding on the subject *Netwok functions for quantum information* Ile de France Région funding on the subject *Quantum Interface for storage of long distance propagating photons* (collaboration with "Institut d'Optique Graduate School").

Objectives

In the domain of classical optical signal processing, diffraction gratings are a basic resource that can be used for a number of devices, ranging from filters to holographic memories. Significant advances can be made, as far as diffractive properties are concerned, if a clever combination of material choice, nonlinear effects and configuration can be found, which has been our main concern for many years.
A new research subject on quantum signal processing for quantum communications applications has started for two years, as in this field also, the need is great for new devices based on nonlinear optics.

Results

The investigation of new intracavity gratings configurations using Gaussian beams [20], gain media [41], thin gratings [52] has given rise to very efficient devices for optical signal processing

applications :
The experimental results obtained with a YAG micro-laser confirmed the theoretical predictions and the advantage of the intracavity gain medium [41]. The diffraction efficiency of the grating is increased by a factor 5000 and the angular selectivity by a factor 20. The developed models enable predictions on various devices from the infinitely thin grating [52] to the thick grating filling the whole cavity that was experimentally tested.
The 2D refractive index gratings, using the band edge resonance of the Bragg mirror to enhance the diffraction properties of the transverse diffraction grating have also been very successful. With the dual independently tunable optical parametric generator developed in our laboratory, a Bragg diffraction regime was observed together with a huge enhancement of the diffraction efficiency in these crystals, in spite of their micrometric size[33]. The simple analytical modeling developed for this kind of gratings can be most useful for the design of new devices [34].

The first achievement concerning quantum signal processing is the implementation of a continuous polarisation entangled photon pairs source at 810 nm, based on spontaneous parametric down-conversion [51]. It was setup for teaching purposes but its performances are comparable to the published results for comparable systems.

The next extraordinary challenge for quantum communication networks is the quantum repeater, including a quantum memory, a full Bell-state analysis and also an entanglement purification facility. The first issue is the compatibility between the long distance carrier photons at 1550 nm, with a bandwidth of 1 nm and the storage systems that operate below 900 nm, with a linewidth of only few hundreds of fm.
In this context, two key elements are a narrowband polarisation entangled photon pairs source and the corresponding wavelength changing interface that will preserve the bandwidth and polarisation of the photons. Nonlinear optics is at the heart of all these functions as spontaneous parametric down conversion will be used for the source, together with very complex filtering, and sum-frequency generation for the interface. An optical parametric oscillator will be setup as a specific narrow-band pumping source for the sum-frequency generation.
With the grants of Region Ile de France and Institut Telecom, the experiments on the quantum interface that will enable the storage of a telecom photon in a solid state quantum memory while preserving its polarization have been started.[181]. The investigation of the compatibility of a propagating qubit with the quantum memory has also led us to the project of designing a new narrow-band polarisation entangled photon pairs source. Future work will be conducted in collaboration with the IQ team of Romain Alléaume (INFRES department of Telecom ParisTech), the Laboratoire Aimé Cotton in Orsay and the LPMC of Nice University within the framework of the three years “eQUANET” ANR project (accepted in 2009). Preliminary experiments show that 20000 photon pairs should be available in the 40 MHz expected bandwidth.

1.3 References

1.3.1 ACL: Articles in ISI-Indexed Journals

[1] A. Aissa El Bey, K. Abed-Meraim, and Y. Grenier. Blind separation of underdetermined convolutive mixtures using their time-frequency representation. *IEEE Transactions on Audio, Speech & Language Processing*, 15(5):1540–1550, July 2007.

[2] A. Aissa El Bey, K. Abed-Meraim, and Y. Grenier. Underdetermined blind audio source separation using modal decomposition. *EURASIP Journal on Audio, Speech & Music Processing*, 2007:1–15, March 2007.

[3] A. Aissa El Bey, K. Abed-Meraim, Y. Grenier, and Y. Hua. A general framework for second order blind separation of stationary colored sources. *Signal Processing*, 88(9):2123–2137, September 2008.

[4] A. Aissa El Bey, N. Linh-Trung, K. Abed-Meraim, A. Belouchrani, and Y. Grenier. Underdetermined

blind separation of nondisjoint sources in the time-frequency domain. *IEEE Transactions on Signal Processing*, 55(3):897–907, March 2007.
[5] M. Alonso, G. Richard, and B. David. Accurate tempo estimation based on harmonic+noise decomposition. *Eurasip Journal on Advances in Signal Processing*, January 2007.
[6] M. Alonso, G. Richard, and B. David. Tempo estimation for audio recordings. *Journal of New Music Research*, 36(1):17–26, March 2007.
[7] R. Badeau and R. Boyer. Fast multilinear singular value decomposition for structured tensors. *SIAM Journal on Matrix Analysis and Applications*, 30(3):1008–1021, September 2008.
[8] R. Badeau, B. David, and G. Richard. Fast approximated power iteration subspace tracking. *IEEE Transactions on Signal Processing*, 53(8 part 1):2931–2941, August 2005.
[9] R. Badeau, B. David, and G. Richard. High resolution spectral analysis of mixtures of complex exponentials modulated by polynomials. *IEEE Transactions on Signal Processing*, 54(4):1341–1350, April 2006.
[10] R. Badeau, B. David, and G. Richard. A new perturbation analysis for signal enumeration in rotational invariance techniques. *IEEE Transactions on Signal Processing*, 54(2):450–458, February 2006.
[11] R. Badeau, B. David, and G. Richard. Cramér-Rao bounds for multiple poles and coefficients of quasipolynomials in colored noise. *IEEE Transactions on Signal Processing*, 56(8):3458–3467, August 2008.
[12] R. Badeau, G. Richard, and B. David. Fast and stable yast algorithm for principal and minor subspace tracking. *IEEE Transactions on Signal Processing*, 56(8):3437–3446, August 2008.
[13] R. Badeau, G. Richard, and B. David. Performance of ESPRIT for estimating mixtures of complex exponentials modulated by polynomials. *IEEE Transactions on Signal Processing*, 56(2):492–504, February 2008.
[14] C. Baras, N. Moreau, and P. Dymarki. Controlling the inaudibility and maximizing the robustness in an audio annotation watermarking system. *IEEE Transactions on Audio, Speech and Language Processing*, 14(5):1772– 1782, September 2006.
[15] C. Berthomier, X. Drouot, M. Herman-Stoïca, P. Berthomier, J. Prado, O. Benoît, J. Mattout, and M. P. D'Ortho. Wake-rem-nrem automatic classification based on a single eeg channel: Epoch by epoch comparison with human sleep scoring in patients. *Journal of Sleep Research*, September 2006.
[16] C. Berthomier, X. Drouot, M. Herman-Stoïca, J. Prado, J. Mattout, and M. P. D'Ortho. Automatic analysis of single-channel sleep eeg: Validation in healthy individuals. *Journal of Sleep research*, 30(11):1587–1595, 2007.
[17] M. Betser, P. Collen, G. Richard, and B. David. Estimation of frequency for am/fm models using the phase vocoder framework. *IEEE Transactions on Signal Processing*, 56(2):505 – 517, February 2008.
[18] D. Bitauld, L. Menez, I. Zaquine, A. Maruani, and R. Frey. Diffraction of Gaussian beams on intracavity bragg gratings. *Journal of the Optical Society of America B*, 22(6):1153–1160, June 2005.
[19] D. Bitauld, I. Zaquine, A. Maruani, and R. Frey. Uniform response high resolution tunable optical filtering using a grating-assisted acousto-optic device. *Optics Express*, 13(17):6438–6444, August 2005.
[20] D. Bitauld, I. Zaquine, A. Maruani, and R. Frey. Numerical analysis of a high resolution fast tunable filter based on an intracavity bragg grating. *Applied Optics*, 46(21):4728–4735, July 2007.
[21] C. Clavel, I. Vasilescu, L. Devillers, G. Richard, and T. Ehrette. Fear-type emotion recognition for future audio-based. *Speech Communication*, 50(6):487–503, June 2008.
[22] O. Derrien, P. Duhamel, M. Charbit, and G. Richard. A new quantization optimization algorithm for the mpeg advanced audio coder using a statistical sub-band model of the quantization noise. *IEEE Transactions on Audio, Speech and Language Processing*, 14(4):1328–1339, July 2006.
[23] O. Derrien and G. Richard. A new model-based algorithm for optimizing the mpeg-aac in ms-stereo. *IEEE Transactions on Audio, Speech and Language Processing*, 16(8):1373–1382, November 2008.
[24] K. Ege, X. Boutillon, and B. David. High-resolution modal analysis. *Journal of Sound and Vibration*, May 2009.
[25] S. Essid, G. Richard, and B. David. Instrument recognition in polyphonic music based on automatic taxonomies. *IEEE Transactions on Audio, Speech, and Language Processing*, 14(1):68–80, January 2006.
[26] S. Essid, G. Richard, and B. David. Musical instrument recognition by pairwise classification strategies. *IEEE Transactions on Audio, Speech, and Language Processing*, 14(4):1401– 1412, July 2006.
[27] C. Févotte, N. Bertin, and J.-L. Durrieu. Nonnegative matrix factorization with the Itakura-Saito divergence. With application to music analysis. *Neural Computation*, 21(3), March 2009.
[28] C. Févotte, B. Torrésani, L. Daudet, and S. J. Godsill. Sparse linear regression with structured priors

and application to denoising of musical audio. *IEEE Trans. Audio, Speech and Language Processing*, 16(1):174–185, January 2008.

[29] O. Gillet, S. Essid, and G. Richard. On the correlation of automatic audio and visual segmentations of music videos. *IEEE Trans. on Circuit and Systems for Video Technology*, March 2007.

[30] O. Gillet and G. Richard. Drum loops retrieval from spoken queries. *Journal of Intelligent Information Systems - Special issue on Intelligent Multimedia Applications*, 24(2/3):159–177, March 2005.

[31] O. Gillet and G. Richard. Transcription and separation of drum signals from polyphonic music. *IEEE Transactions on Audio, Speech and Language Processing*, 16(3):529 – 540, March 2008.

[32] M. Guillaume and Y. Grenier. Sound field analysis based on analytical beamforming. *EURASIP Journal on Advances in Signal Processing*, 2007, August 2007.

[33] Q. He, I. Zaquine, R. Frey, R. Andre, and G. Roosen. Efficient bragg diffraction in thin semiconductor 2d gratings. *Optics Letters*, 33(23):2868–2870, December 2008.

[34] Q. He, I. Zaquine, R. Frey, and G. Roosen. Bragg diffraction in thin 2d refractive index modulated semiconductor samples. *Journal of the Optical Society of America B*, 26(3):390–396, March 2009.

[35] Q. He, I. Zaquine, A. Maruani, and R. Frey. Band edge induced bragg diffraction in 2d photonic crystals. *Optics Letters*, 31(9):1184–1186, May 2006.

[36] Th. Hélie and D. Matignon. Diffusive reprentations for the analysis and simulation of flared acoustic pipes with visco-thermal losses. *Mathematical Models and Methods in Applied Sciences (M3AS)*, 16(4):503–536, April 2006.

[37] Th. Hélie and D. Matignon. Representations with poles and cuts for the time-domain simulation of fractional systems and irrational transfer functions. *Signal Processing (SP)*, 86(10):2516–2528, October 2006.

[38] C. Joder, S. Essid, and G. Richard. Temporal integration for audio classification with application to musical instrument classification. *IEEE Transaction on Audio, Speech and Language Processing*, 17(1):174–186, January 2009.

[39] J. Kergomard, V. Debut, and D. Matignon. Resonance modes in a 1d medium with two purely absorbing boundaries: calculation methods, orthogonality and completeness. *Journal of the Acoustical Society of America (JASA)*, 119(3):1356–1367, March 2006.

[40] P. Leveau, E. Vincent, G. Richard, and L. Daudet. Instrument-specific harmonic atoms for mid-level music representation. *IEEE Transactions on Audio, Speech and Language Processing*, January 2008.

[41] A. Moreau, Q. He, I. Zaquine, A. Maruani, and R. Frey. Intracavity gain gratings. *Optics Letters*, 32(3):208–210, February 2007.

[42] A. Moreau, I. Zaquine, A. Maruani, and R. Frey. Efficient Bragg-like operation of intracavity low efficiency plane gratings. *Journal of the Optical Society of America B*, 22(11):2289–2294, November 2005.

[43] A. Ozerov and C. Févotte. Multichannel nonnegative matrix factorization in convolutive mixtures for audio source separation. *IEEE Trans. Audio, Speech and Language Processing*, 2010.

[44] E. Ravelli, G. Richard, and L. Daudet. Union of mdct bases for audio coding. *IEEE Transactions on Audio, Speech and Language Processing*, 16(8):1361–1372, November 2008.

[45] E. Ravelli, G. Richard, and L. Daudet. Audio signal representations for indexing in the transform domain. *IEEE Transactions on Audio, Speech and Language Processing*, December 2009.

1.3.2 ACLN: Articles in Other Refereed Journals

[46] C. Clavel, I. Vasilescu, G. Richard, and L. Devillers. Du corpus émotionnel au système de détection : le point de vue applicatif de la surveillance dans les lieux publics. *Revue en Intelligence Artificielle (RIA)*, 20(4-5):529–551, September 2006.

[47] Y. Grenier. Interfaces audio aes/ebu. In *Editions Techniques de l'Ingénieur*. Editions Techniques de l'Ingénieur, 2005.

[48] Th. Hélie, D. Matignon, and R. Mignot. Criterion design for optimizing low-cost approximations of infinite-dimensional systems: towards efficient real-time simulation. *International Journal of Tomography and Statistics*, 7(7):13–18, September 2007.

[49] H. Kim, K. Honda, and S. Maeda. Stroboscopic cine-mri study on the phasing between the tongue and the larynx in korean three-way phonation contrast. *Journal of Phonetics*, 33(1):1–26, January 2005.

[50] D. Matignon and Ch. Prieur. Asymptotic stability of linear conservative systems when coupled with diffusive systems. *European Series in Applied and Industrial Mathematics: Control, Optimization and Calculus of Variations (ESAIM:COCV)*, 11(3):487–507, July 2005.

[51] Y. Menesguen, J. L. Smirr, G. Pillet, R. Alleaume, A. Maruani, I. Zaquine, R. Frey, and L. Jacubowiez. Source de photons intriqués en polarisation : travaux pratiques de physique quantique. *Bulletin de l'Union des Physiciens*, 102:61–80, November 2008.
[52] A. Moreau, I. Zaquine, A. Maruani, and R. Frey. Réseaux minces de diffraction en régime de bragg. *Journal de Physique IV*, 135:239, October 2006.
[53] J. Prado. Introduction à matlab. In *Techniques de l'Ingénieur*, chapter dossier AF1450, Vol papier n^{o}AFM3. 2005.

1.3.3 ACTI: Articles in Proceedings of International Conferences

[54] A. Aissa El Bey, K. Abed-Meraim, and Y. Grenier. Blind separation of audio sources convolutive mixtures using parametric decomposition. In *IWAENC'05*, volume 1, pages 161–164, Eindhoven (Pays-bas), September 2005.
[55] A. Aissa El Bey, K. Abed-Meraim, and Y. Grenier. Blind separation of audio sources using modal decomposition. In *ISSPA'05*, volume 2, pages 451–454, Sydney (Australie), August 2005.
[56] A. Aissa El Bey, K. Abed-Meraim, and Y. Grenier. Underdetermined blind source separation of audio sources in time-frequency domain. In *SPARS'05*, volume 1, pages 115–118, Rennes, France, November 2005.
[57] A. Aissa El Bey, K. Abed-Meraim, and Y. Grenier. Iterative blind source separation by decorrelation: algorithm and performance analysis. In *14th European signal processing conference (EUSIPCO)*, Florence, Italie, September 2006.
[58] A. Aissa El Bey, K. Abed-Meraim, and Y. Grenier. On the identifiability testing in blind source separation using resampling technique. In *6th International Conference on Independent Component Analysis and Blind Source Separation*, number LCNS 3889, pages 755–764, Charleston, SC, USA, March 2006.
[59] A. Aissa El Bey, K. Abed-Meraim, and Y. Grenier. Blind audio source separation using sparsity based criterion for convolutive mixture case. In *7th International Conference on Independent Component Analysis and Blind Source Separation*, London, UK, September 2007.
[60] A. Aissa El Bey, K. Abed-Meraim, and Y. Grenier. Underdetermined audio source separation using fast parametric decomposition. In *ISSPA'07*, Sharjah (United Arab Emirates), February 2007.
[61] A. Aissa El Bey, K. Abed-Meraim, and Y. Grenier. Underdetermined blind separation of audio sources from the time-frequency representation of their convolutive mixtures. In *ICASSP'07*, volume 1, pages 153–156, Hawaii (USA), April 2007.
[62] A. Aissa El Bey, H. Bousbia-Salah, K. Abed-Meraim, and Y. Grenier. Audio source separation using sparsity. In *IWAENC'06*, Paris, France, September 2006.
[63] M. Alonso, G. Richard, and B. David. Extracting note onsets from musical recordings. In *IEEE-ICME*, Amsterdam, NL, July 2005.
[64] J. Altuna, B. Mulgrew, R. Badeau, and V. Atxa. A fast adaptive method for subspace based blind channel estimation. In *ICASSP'06*, volume IV, pages 1121–1124, Toulouse, France, May 2006.
[65] R. Badeau and B. David. Weighted Maximum Likelihood Autoregressive and Moving Average Spectrum Modeling. In *ICASSP'08*, pages 3761–3764, Las Vegas, Nevada, USA, April 2008.
[66] R. Badeau, B. David, and G. Richard. Yet another subspace tracker. In *ICASSP'05*, volume 4, pages 329–332, Philadelphie, Pennsylvanie, USA, March 2005.
[67] R. Badeau, B. David, and G. Richard. YAST algorithm for minor subspace tracking. In *ICASSP'06*, volume III, pages 552–555, Toulouse, France, May 2006.
[68] R. Badeau, B. David, and G. Richard. Conjugate gradient algorithms for minor subspace analysis. In *ICASSP'07*, volume 3, pages 1013–1016, Honolulu, Hawaii, USA, April 2007.
[69] R. Badeau, V. Emiya, and B. David. Expectation-maximization algorithm for multi-pitch estimation and separation of overlapping harmonic spectra. In *ICASSP'09*, pages 3073–3076, Taipei, Taiwan, April 2009.
[70] R. Badeau, G. Richard, and B. David. Fast adaptive esprit algorithm. In *SSP'05*, Bordeaux, France, July 2005.
[71] W. Bailer, E. Dumont, S. Essid, and B. Mérialdo. A collaborative approach to automatic rushes video summarization. In *IEEE ICIP Workshop on Multimedia Information Retrieval: New Trends and Challenges*, October 2008.
[72] C. Baras and N. Moreau. An audio spread-spectrum data hiding system with an informed embedding strategy adapted to a Wiener filtering based receiver. In *IEEE International Conference on Multimedia and Exposition (ICME)*, Amsterdam, NL, July 2005.

[73] C. Baras, N. Moreau, and P. Dymarki. Comparative study of two informed embedding strategies for audio spread-spectrum data hiding systems. In *EUSIPCO*, Antalya, Turquie, September 2005.

[74] C. Berthomier, X. Drouot, M. Herman-Stoïca, J. Prado, J. Mattout, and M. P. D'Ortho. A wake-rem-nrem automatic analysis using a single eeg channel: Epoch by epoch comparison with human sleep scoring in healthy subjects. In *First Congress of the World Association of Sleep Medicine (WASM)*, Berlin, Allemagne, October 2005.

[75] C. Berthomier, X. Drouot, M. Herman-Stoïca, J. Prado, J. Mattout, and M. P. D'Ortho. Single channel based brain monitoring: Sleep/wakefulness classification. In *International conference on Monitoring sleep and sleepiness - from physiology to new sensors*, Bâle, Suisse, May 2006.

[76] C. Berthomier, X. Drouot, M. Herman-Stoïca, J. Prado, J. Mattout, and M. P. D'Ortho. Real-time automatic measurement of recorded sleep time. In *American College of Chest Physicians (ACCP) congress : CHEST*, number 694S, Chicago, US, October 2007.

[77] C. Berthomier, A. Muzet, P. Berthomier, J. Prado, and J. Mattout. Real-time automatic measurement of drowsiness based on a single eeg channel. In *European Sleep Research Society*, Glasgow Scotland, September 2008.

[78] N. Bertin, R. Badeau, and G. Richard. Blind signal decompositions for automatic transcription of polyphonic music: NMF and K-SVD on the benchmark. In *ICASSP'07*, volume 1, pages 65–68, Honolulu, Hawaii, USA, April 2007.

[79] N. Bertin, C. Févotte, and R. Badeau. A tempering approach for Itakura-Saito non-negative matrix factorization. With application to music transcription. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'09)*, pages 1545–1548, Taipei, Taiwan, April 2009.

[80] M. Betser, P. Collen, and G. Richard. Frequency estimation based on adjacent dft bins. In *EUSIPCO-2006*, Florence, Italie, September 2006.

[81] M. Betser, P. Collen, G. Richard, and B. David. Review and discussion on classical stft-based frequency estimators. In *International Convention of the Audio Engineering Society (AES)*, Paris, France, May 2006.

[82] D. Bitauld, I. Zaquine, A. Maruani, and R. Frey. A fast tunable high resolution filter. In *Conference on lasers and electro-optics*, number CI-4-TUE, Munich (Allemagne), June 2005.

[83] D. Bitauld, I. Zaquine, A. Maruani, and R. Frey. Grating-assisted acousto-optic filtering. In *Photonics*, page 7, Prague, République Tchèque, June 2005.

[84] R. Blouet, G. Rapaport, I. Cohen, and C. Févotte. Evaluation of several strategies for single sensor speech/music separation. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'08)*, Las Vegas, USA, April 2008.

[85] R. Boyer, K. Abed-Meraim, and L. De Lathawer. Delayed exponential fitting by best tensor rank-(r1,r2,r3) approximation. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, May 2005.

[86] R. Boyer and R. Badeau. Adaptive multilinear SVD for structured tensors. In *ICASSP'06*, volume III, pages 880–883, Toulouse, France, May 2006.

[87] A. Capra, S. Fontana, F. Adriaensen, A. Farina, and Y. Grenier. Listening tests of the localization performance of stereodipole and ambisonic systems. In *123rd Convention of the Audio Engineering Society*, New York, USA, October 2007.

[88] C. Clavel, L. Devillers, G. Richard, I. Vasilescu, and T. Ehrette. Abnormal situation detection and analysis through fear-type acoustic manifestations. In *ICASSP'07*, Honolulu, Hawai, April 2007.

[89] C. Clavel, T. Ehrette, and G. Richard. Events detection for an audio-based surveillance system. In *IEEE-ICME*, Amsterdam, July 2005.

[90] C. Clavel, I. Vasilescu, L. Devillers, G. Richard, T. Ehrette, and C. Sedogbo. The safe corpus: illustrating extreme emotions in dynamic situations. In *LREC Workshop Corpora and Emotion*, Gênes, Italie, May 2006.

[91] C. Clavel, I. Vasilescu, G. Richard, and L. Devillers. Fear-type emotions in the safe corpus : annotation issues. In *LREC*, Gênes, Italie, May 2006.

[92] C. Clavel, I. Vasilescu, G. Richard, and L. Devillers. Voiced and unvoiced content of fear-type emotions in the safe corpus. In *Speech Prosody 2006*, Dresden, Germany, May 2006.

[93] G. Cornuz, E. Ravelli, P. Leveau, and L. Daudet. Object Coding of Harmonic Sounds using Sparse and Structured Representations. In *10th Int. Conference on Digital Audio Effects (DAFx-07)*, Bordeaux, France, September 2007.

[94] B. David and R. Badeau. Fast sequential LS estimation for sinusoidal modeling and decomposition of audio signals. In *WASPAA 2007*, pages 211–214, New Paltz, New York, USA, October 2007.

[95] B. David, R. Badeau, and G. Richard. HRHATRAC algorithm for spectral line tracking of musical

signals. In *ICASSP'06*, volume III, pages 45–48, Toulouse, France, May 2006.

[96] B. David, V. Emiya, R. Badeau, and Y. Grenier. Harmonic plus noise decomposition: Time-frequency reassignment versus a subspace-based method. In *120th Convention of the Audio Engineering Society*, Paris, France, May 2006.

[97] J. Dellière, A. Maruani, H. Maître, and P. Benjamin. A full electromagnetic SAR image simulator for urban structures. In *4th IEEE-GRSS - ISPRS workshop - URBAN 2007*, Paris (France), April 2007.

[98] J. Dellière, A. Maruani, H. Maître, P. Benjamin, and J. P. Piau. A full electromagnetic SAR simulator for urban structures. In *Physics in Signal and Image Processing, PSIP'07*, Mulhouse, January 2007.

[99] J.-F. Deü and D. Matignon. A coupled Nemark-diffusive scheme for fractionally damped oscillators. In *8th International Conference on Mathematical and Numerical Aspects of Waves*, pages 526–528, Reading, United Kingdom, July 2007.

[100] E. Dumont, B. Merialdo, S. Essid, W. Bailer, D. Byrne, H. Bredin, N. E. O'Connor, G. J. F. Jones, M. Haller, A. Krutz, Th. Sikora, and T. Piatrik. A collaborative approach to video summarization. In *SAMT 2008, 3rd International Conference on Semantic and Digital Media Technologies*, Koblenz, Germany, December 2008.

[101] E. Dumont, B. Merialdo, S. Essid, W. Bailer, H. Rehatschek, D. Byrne, H. Bredin, N. E. O'Connor, G. J. F. Jones, A. F. Smeatonand, M. Haller, A. Krutz, Th. Sikora, and T. Piatrik. Rushes video summarization using a collaborative approach. In *TRECVID 2008, ACM International Conference on Multimedia Information Retrieval 2008*, Vancouver, BC, Canada, November 2008.

[102] J.-L. Durrieu, A. Ozerov, C. Févotte, G. Richard, and B. David. Main instrument separation from stereophonic audio signals using a source/filter model. In *European Signal Processing Conference (EUSIPCO)*, Glasgow,Scotland, August 2009.

[103] J.-L. Durrieu, G. Richard, and B. David. Singer melody extraction in polyphonic signals using source separation methods. In *ICASSP'08*, Las Vegas, Nevada, USA, April 2008.

[104] J.-L. Durrieu, G. Richard, and B. David. An iterative approach to monaural musical mixture de-soloing. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, Taiwan, April 2009.

[105] V. Emiya, R. Badeau, and B. David. Multipitch estimation of inharmonic sounds in colored noise. In *10th Int. Conf. on Digital Audio Effects (DAFx-07)*, pages 93–98, Bordeaux, France, September 2007.

[106] V. Emiya, R. Badeau, and B. David. Automatic transcription of piano music based on HMM tracking of jointly-estimated pitches. In *EUSIPCO 2008*, Lausanne, Switzerland, August 2008.

[107] V. Emiya, B. David, and R. Badeau. A parametric method for pitch estimation of piano tones. In *ICASSP'07*, volume 1, pages 249–252, Honolulu, Hawaii, USA, April 2007.

[108] V. Emiya, B. David, and V. Gibiat. Two representation tools to analyse non-stationary sounds in a perceptual context. In *Forum Acusticum 2005*, Budapest, Hongrie, August 2005.

[109] S. Essid, G. Richard, and B. David. Instrument recognition in polyphonic music. In *ICASSP'05*, Philadelphie, US, March 2005.

[110] S. Essid, G. Richard, and B. David. Hierarchical classification of musical instruments on solo recordings. In *ICASSP'06*, Toulouse, France, May 2006.

[111] C. Févotte and F. Theis. Pivot selection strategies in jacobi joint block-diagonalization. In *7th International Conference on Independent Component Analysis and Signal Separation (ICA'07)*, Londres (United-Kingdom), September 2007.

[112] S. Fontana, A. Farina, and Y. Grenier. Binaural for popular music: a case of study. In *13th International Conference on Auditory Display*, Montréal, Canada, June 2007.

[113] S. Fontana, Y. Grenier, and A. Farina. A system for head related impulse responses rapid measurement and direct customization. In *120th Convention AES*, Paris, France, October 2006.

[114] O. Gillet and G. Richard. Automatic transcription of drum sequences using audiovisual features. In *ICASSP'05*, Philadelphie, US, March 2005.

[115] O. Gillet and G. Richard. Drum track transcription of polyphonic music signals using noise subspace projection. In *6th International Conference on Music Information Retrieval, ISMIR 2005*, London, UK, September 2005.

[116] O. Gillet and G. Richard. Extraction and remixing of drum tracks from polyphonic music signals. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, WASPAA'05*, New Paltz, USA, October 2005.

[117] O. Gillet and G. Richard. Indexing and querying drum loops databases. In *CBMI*, Riga, Lettonie, June 2005.

[118] O. Gillet and G. Richard. Comparing audio and video segmentations for music videos indexing. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP'06*, Toulouse,

France, May 2006.

[119] O. Gillet and G. Richard. Enst-drums: an extensive audio-visual database for drum signals processing. In *7th International Conference on Music Information Retrieval, ISMIR 2006*, Victoria, Canada, October 2006.

[120] O. Gillet and G. Richard. Supervised and unsupervised sequence modelling for drum transcription. In *International Congress on Music Information Retrieval (ISMIR)*, Vienne, Autriche, September 2007.

[121] S. J. Godsill, A. T. Cemgil, C. Févotte, and P. J. Wolfe. Bayesian computational methods for sparse audio and music processing. In *Proc. 15th European Signal Processing Conference (EUSIPCO'07)*, Poznan (Poland), September 2007.

[122] M. Guillaume and Y. Grenier. Experimental 3d sound field analysis with a microphone array. In *28th International Conference of the AES*, Pitea, Sweden, July 2006.

[123] M. Guillaume and Y. Grenier. Sound field analysis based on generalized prolate spheroidal wave sequences. In *120th Convention of the Audio Engineering Socitety*, Paris, FRANCE, May 2006.

[124] M. Guillaume and Y. Grenier. Sound field analysis with a two-dimensional microphone array. In *ICASSP*, volume V, pages 321–324, Toulouse, France, May 2006.

[125] M. Guillaume, Y. Grenier, and G. Richard. Iterative algorithms for multichannel equalization in sound reproduction systems. In *ICASSP'05*, Philadelphie, US, March 2005.

[126] H. Haddar, J.-R. Li, and D. Matignon. Efficient solution of a wave equation with fractional order dissipative terms. In *8th International Conference on Mathematical and Numerical Aspects of Waves*, pages 529–531, Reading, United Kingdom, July 2007.

[127] Z. Harchaoui, F. Vallet, A. Lung-Yut-Fong, and O. Cappé. A regularized kernel-based approach to unsupervised audio segmentation. In *ICASSP 2009*, pages 1665–1668, Taiwan, April 2009.

[128] Q. He, I. Zaquine, R. Andre, G. Roosen, and R. Frey. Bragg diffraction regime in thin semiconductor 2d refractive index gratings. In *PR 09*, Bad Honnef, Allemagne, June 2009.

[129] Q. He, I. Zaquine, A. Maruani, S. Massenot, R. Chevallier, and R. Frey. Diffraction enhancement in 2d photonic crystals. In *European Optical Society Annual Meeting*, volume TOM3, page 193, Paris France, October 2006.

[130] Th. Hélie, D. Matignon, and R. Mignot. Criterion design for optimizing low-cost approximations of infinite-dimensional systems: towards efficient real-time simulation. In *IFAC workshop on Control Applications of Optimisation (CAO'06)*, pages 368–373, Cachan, France, April 2006.

[131] C. Hory and W. J. Christmas. Cepstral features for classification of an impulse response with varying sample size dataset. In *EUSIPCO 2007*, pages 1546–1550, Poznan, Pologne, September 2007.

[132] C. Joder, S. Essid, and G. Richard. Alignment kernels for audio classification with application to music instrument recognition. In *EUSIPCO 2008*, Lausanne, Suisse, August 2008.

[133] J. Kergomard, D. Matignon, and V. Debut. Waves in a 1d-medium with two resistive terminations: completeness of the modes. In *WAVES'05: 7th Int. Conf. on Mathematical and Numerical Aspects of Wave Propagation*, pages 273–275, Providence (Rhode Island, USA), June 2005.

[134] S. Krstulovic, R. Gribonval, P. Leveau, and L. Daudet. A comparison of two extensions of the matching pursuit algorithm for the harmonic decomposition of sounds. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, (WASPAA05)*, Mohonk (NY) - USA, October 2005.

[135] M. Lardeur, S. Essid, G. Richard, M. Haller, and T. Sikora. Incorporating prior knowledge on the digital media creation process into audio classifiers. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, Taiwan, April 2009.

[136] J.-L. Le Carrou, F. Gautier, and R. Badeau. Theoretical and experimental investigations of harp's sympathetic modes. In *ICA 2007*, Madrid, Spain, September 2007.

[137] P. Leveau and L. Daudet. Multi-resolution partial tracking with modified matching pursuit. In *14th European Signal Processing Conference (EUSIPCO)*, Florence (Italy), September 2006.

[138] P. Leveau, S. Essid, G. Richard, L. Daudet, and B. David. On the usefulness of differentiated transient/steady-state processing in machine recognition of musical instruments. In *AES convention*, Barcelone, May 2005.

[139] P. Leveau, D. Sodoyer, and L. Daudet. Automatic Instrument Recognition in a Polyphonic Mixture using Sparse Representations. In *8th Int. Conf. on Music Information Retrieval (ISMIR 2007)*, Vienne, Autriche, September 2007.

[140] P. Leveau, E. Vincent, G. Richard, and L. Daudet. Mid-level sparse representations for timbre identification: design of an instrument-specific harmonic dictionary. In *1st Workshop on Learning the Semantics of Audio Signals*, Athènes, Grèce, December 2006.

[141] N. Linh-Trung, A. Aissa El Bey, K. Abed-Meraim, and A. Belouchrani. Underdetermined blind source separation of non-disjoint nonstationary sources in time-frequency domain. In *ISSPA'05*, volume 1,

pages 46–49, Sydney (Australie), August 2005.

[142] S. Maeda and M. Toda. Quantal aspects of non anterior sibilant fricatives: A simulation study. In *7th International Seminar on Speech Production (ISSP)*, pages 573–580, July 2006.

[143] D. Matignon. Asymptotic stability of the Webster-Lokshin model. In *Mathematical Theory of Networks and Systems*, page CDROM 11p, Kyoto, Japan, July 2006.

[144] D. Matignon. Diffusive representation for fractional Laplacian and other non-causal pseudo-differential operato rs. In *IFAC workshop on Control of Distributed Parameter Systems (CDPS'07)*, pages 19–20, Namur, Belgique, July 2007.

[145] K. Mcguinness, O. Gillet, N. O'Connor, and G. Richard. Visual analysis for drum sequence transcription. In *European Signal Processing Conference (Eusipco)*, Poznan, Pologne, September 2007.

[146] R. Mignot, Th. Hélie, and D. Matignon. Waveguide modeling of lossy flared acoustic pipes: derivation of a Kelly-Lochbaum structure for real-time simulations. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, page 4, New Paltz, USA, October 2007.

[147] A. Moreau, Q. He, I. Zaquine, A. Maruani, and R. Frey. Intracavity bragg diffraction in microlasers. In *European Optical Society Annual Meeting*, volume TOM6, page 45, Paris France, October 2006.

[148] A. Moreau, Q. He, I. Zaquine, A. Maruani, and R. Frey. Gain grating in a nd:yvo4 microlaser. In *Conference on Lasers and Electrooptics*, Baltimore, May 2007.

[149] A. Moreau, I. Zaquine, A. Maruani, and R. Frey. Diffraction regime of an intracavity thin grating. In *Conference on Lasers and Electro-Otpics*, number CK-4-WED, Munich (Allemagne), June 2005.

[150] A. Moreau, I. Zaquine, A. Maruani, and R. Frey. Intracavity thin diffraction gratings. In *Photonics*, page 83, Prague, République Tchèque, June 2005.

[151] I. Potamitis and A. Ozerov. Single channel source separation using static and dynamic features in the power domain. In *EUSIPCO, 16th European Signal Processing Conference*, Laussane, Switzerland, August 2008.

[152] M. Ramona, G. Richard, and B. David. Vocal detection in music with support vector machines. In *ICASSP'08*, Las Vegas, USA, April 2008.

[153] E. Ravelli, G. Richard, and L. Daudet. Extending transform coding to very lowbirates using overcomplete dictionaries. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, US, October 2007.

[154] E. Ravelli, G. Richard, and L. Daudet. Fast mir in a sparse transform domain. In *ISMIR*, Philadelphie, USA, September 2008.

[155] E. Ravelli, G. Richard, and L. Daudet. Matching pursuit in adaptive dictionaries for scalable audio coding. In *EUSCIPCO*, Lausanne, Suisse, September 2008.

[156] G. Richard, P. Leveau, L. Daudet, S. Essid, and B. David. Towards polyphonic musical instrument recognition. In *International Congress on Acoustics (ICA)*, Madrid, September 2007.

[157] G. Richard, M. Ramona, and S. Essid. Combined supervised and unsupervised approaches for automatic segmentation of radiophonic audio streams. In *ICASSP'07*, Honolulu, Hawai, April 2007.

[158] W. Souidene, A. Aissa El Bey, K. Abed-Meraim, and A. Beghdadi. Blind image separation using sparse representation. In *14th International Conference on Image Processing ICIP*, San Antonio, Texas, USA, September 2007.

[159] V. Y. F. Tan and C. Févotte. Automatic relevance determination in nonnegative matrix factorization. In *Workshop on Signal Processing with Adaptative Sparse Structured Representations (SPARS'09)*, St-Malo, France, April 2009.

[160] M. Toda, S. Maeda, M.Aron, and M.0. Berger. Modeling subject-specific formant transition patterns in /asha/ sequences. In *8th International Seminar on Speech Production (ISSP)*, pages 357–360, July 2008.

[161] K. Trabelsi, T. Hélie, and D. Matignon. Time-domain simulation of functions and dynamical systems of Bessel type. In *8th International Conference on Mathematical and Numerical Aspects of Waves*, pages 547–549, Reading, United Kingdom, July 2007.

[162] I. Venturini. Oracle attacks and covert channels. In *International Workshop on Digital Watermarking, IWDW 2005*, Siena, Italy, September 2005.

[163] I. Venturini. Reality Preserving Fractional Discrete Cosine Transforms. In *Electronic imaging and the Visual Arts 2006*, Florence, Italy, April 2006.

[164] E. Vincent, N. Bertin, and R. Badeau. Harmonic and inharmonic nonnegative matrix factorization for polyphonic pitch transcription. In *ICASSP'08*, pages 109–112, Las Vegas, Nevada, USA, April 2008.

[165] S. Wegener, M. Haller, J.-J. Burred, T. Sikora, S. Essid, and G. Richard. On the robustness of audio features for musical instrument classification. In *EUSCIPCO*, Lausanne, Suisse, September 2008.

[166] P. Wilkins, T. Adamek, D. Byrne, G. Jones, H. Lee, G. Keenan, K. Mcguinness, N. E. O'Connor, A. F.

Smeaton, A. Amin, Z. Obrenovic, R. Benmokhtar, E. Galmar, B. Huet, S. Essid, R. Landais, F. Vallet, G. Th. Papadopoulos, S. Vrochidis, V. Mezaris, I. Kompatsiaris, E. Spyrou, Y. Avrithis, R. Morzinger, P. Schallauer, W. Bailer, T. Piatrik, K. Chandramouli, E. Izquierdo, M. Haller, L. Goldmann, A. Samour, A. Cobet, T. Sikora, and P. Praks. K-space at TRECVid 2007. In *TRECVID 2007*, November 2007.

[167] M. Yeou and S. Maedaand K.S.Honda. Laryngeal activity in the production of consonants clusters and geminates in moroccan arabic. In *8th International Seminar on Speech Production (ISSP)*, pages 249–252, July 2008.

1.3.4 ACTN: Articles in Proceedings of French Conferences

[168] A. Aissa El Bey, K. Abed-Meraim, and Y. Grenier. Séparation aveugle sous-déterminée de sources audio par la méthode EMD (Empirical Mode Decomposition). In *20e Colloque GRETSI sur le traitement du signal et des images*, volume 2, pages 1233–1236, Louvain-La-Neuve, Belgique, September 2005.

[169] A. Aissa El Bey, K. Abed-Meraim, and Y. Grenier. Séparation aveugle sous-déterminée de sources en utilisant la décomposition en paquet d'ondelettes. In *21e Colloque GRETSI sur le traitement du signal et des images*, Troyes, France, September 2007.

[170] G. Bailly, C. Baras, P. Bas, S. Baudry, R. Brun, D. Beautemps, F. Davoine, F. Elisei, G. Gibert, D. Grison, J.-P. Léoni, J. Liénard, N. Moreau, and P. Nguyen. Artus : Calcul et tatouage audiovisuel des mouvements d'un personnage animé virtuel pour l'accessibilité d'émissions télévisuelles aux spectateurs sourds comprenant la langue française parlée complétée. In *Handicap*, pages 265–270, Paris, June 2006.

[171] C. Baras and N. Moreau. Modulation cdma informée dans un système de tatouage audio. In *CORESA*, Rennes, France, November 2005.

[172] C. Baras, N. Moreau, and B. Zayen. Mécanisme de synchronisation en tatouage audio pour des perturbations désynchronisantes à forte dérive. In *GRETSI*, pages 1205–1208, Louvain-La-Neuve, Belgique, September 2005.

[173] C. Berthomier, X. Drouot, M. Herman-Stoïca, J. Prado, J. Mattout, and M. P. D'Ortho. Analyse automatique du sommeil à partir d'une dérivation eeg : Comparaison entre hypnogrammes automatique et visuel chez le sujet sain. In *XXème Congrès de la Société Française de Recherche sur le Sommeil (SFRS)*, Lyon, November 2005.

[174] C. Berthomier, X. Drouot, M. Herman-Stoïca, J. Prado, J. Mattout, and M. P. D'Ortho. Analyse automatique du sommeil à partir dune unique dérivation eeg : Comparaison des analyses automatique et visuelle chez le patient. In *XXIème Congrès de la Société Française de Recherche et Médecine sur le Sommeil (SFRMS)*, Albi, France, November 2006.

[175] C. Berthomier, X. Drouot, M. Herman-Stoïca, J. Prado, J. Mattout, and M. P. D'Ortho. Sleep automatic analysis using a single electroencephalogram channel in healthy subjects. In *Congrès de Physiologie, de Pharmacologie et Thérapeutique (P2T)*, Lyon, April 2006.

[176] D. Bitauld, I. Zaquine, A. Maruani, and R. Frey. Filtrage frequentiel optique par un dispositif acousto-optique associe a un reseau de diffraction. In *COLOQ 9*, page 43, Dijon, France, September 2005.

[177] K. Ege, X. Boutillon, B. David, and J. Lozada. Analyse modale sans transformée de fourier. In *Congrès Français de Mécanique*, Grenoble, August 2007.

[178] Q. He, I. Zaquine, A. Maruani, and R. Frey. Diffraction de bragg dans des cristaux photoniques 2d. In *Coloq*, Grenoble, July 2007.

[179] J.-L. Le Carrou, F. Gautier, and R. Badeau. Analyse des modes de cordes couplées dune harpe par une méthode à haute résolution. In *8ème Congrès Français d'Acoustique CFA'06*, Tours, France, April 2006.

[180] J. Lozada, X. Boutillon, and B. David. Modes propres de vibration : comment oublier fourier. In *Congrès Français de Mécanique*, Tours, France, April 2006.

[181] Y. Menesguen, J. L. Smirr, G. Pillet, C. Bourdarias, R. Alleaume, A. Maruani, I. Zaquine, and R. Frey. Interface de changement de longueur donde conservant lintrication en polarisation. In *COLOQ*, Grenoble, July 2007.

[182] A. Moreau, I. Zaquine, A. Maruani, and R. Frey. Réseaux minces de diffraction en régime de bragg. In *COLOQ 9*, page 189, Dijon France, September 2005.

[183] M. Ramona and G. Richard. Segmentation parole/musique par machines à vecteurs de support. In *Journées d'Etudes sur la Parole (JEP'08)*, Avignon, France, June 2008.

1.3.5 COM: Talks in Conferences Which Do Not Publish Proceedings

[184] R. Badeau and B. David. Adaptive subspace methods for high resolution analysis of music signals. In *Acoustics'08*, Paris, France, July 2008.
[185] N. Bertin and R. Badeau. Initialization, distances and local minima in audio applications of the non-negative matrix factorization. In *Acoustics'08*, Paris, France, July 2008.
[186] B. David and R. Badeau. Towards an adaptive subspace-based representation of musical spectral content. In *Acoustics'08*, Paris, France, July 2008.
[187] O. Derrien, G. Richard, and R. Badeau. Damped sinusoids and subspace based approach for lossy audio coding. In *Acoustics'08*, Paris, France, July 2008.
[188] J.-L. Durrieu, G. Richard, and B. David. Single sensor singer/music separation using a source/filter model of the singer voice. In *Acoustics*, Paris, France, July 2008.
[189] J.-L. Durrieu and J. Weil. Automatic beat-synchronous generation of music lead sheets. In *Kspace PhD Jamboree*, Paris, France, July 2008.
[190] F. Vallet, G. Richard, S. Essid, and J. Carrive. Detecting artist performances in a tv show. In *Kspace PhD Jamboree*, Paris, France, July 2008.

1.3.6 OS: Books and Book Chapters

[191] C. Baras, N. Moreau, and T. Dutoit. *How could music contain hidden information*, chapter 7, pages 223 – 264. 2009.
[192] R. Benmokhtar, B. Huet, G. Richard, T. Declerck, and S. Essid. *Feature Extraction for Multimedia Analysis*, chapter 4. Wiley, 2009.
[193] C. Clavel and G. Richard. *Reconnaissance acoustique des émotions,*, chapter 5. Hermès, 2009.
[194] T. Dutoit and N. Moreau. *How is sound processed in an MP3 player*, chapter 3, pages 65–101. 2009.
[195] T. Dutoit, N. Moreau, and P. Kroon. *How is speech processed in a cell phone conversation*, chapter 1, pages 1–31. 2009.
[196] S. Essid, M. Campedel, G. Richard, T. Piatrik, R. Benmokhtar, and B. Huet. *Machine Learning Techniques for Multimedia Analysis*, chapter 5. Wiley, 2009.
[197] C. Févotte. *Bayesian audio source separation*, chapter 11, pages 305–335. Springer, 2007.
[198] N. Moreau. *Codage audio et normes*. Vuibert, 2006.
[199] G. Richard. *Audio Indexing*. Information Science Reference - IGI Global, 2008.
[200] M. Toda, S. Maeda, and K. Honda. *Formant-cavity affiliation in sibilant fricatives*. Fuchs, S., Zygis, M., Toda, M., and Shadle Ch. (Eds.). Mouton de Gruyter., 2009.