Part IV

Signal and Image Processing
The research topics covered by the Signal and Image Processing department at TELECOM Paris-Tech are: the study of image processing in its various formats, digital, optical... for different applications like medical imaging, remote sensing, fine arts..., the study of speech, music and sound.

After its reorganization at the beginning of 2007, the department is now organized into four groups:

- "Statistics and applications" - STA - is a group that is devoted to the applications of statistics and probability to the field of information processing. The research area covers a wide spectrum from the development of new techniques and new algorithms to various applications. The activities of the group comprises the following topics: statistical learning, independent data and complex random systems, methods and algorithms for cosmological data analysis, Markov Chain Monte-Carlo techniques, sequential Monte-Carlo techniques (particle filters), array processing, geolocalization, models estimation.

- "Image Processing and Interpretation" - TII - has, as its main purpose, the development of methodologies and theoretical tools for image processing, scene analysis and 3D objects. This implies global treatment of complex image processing problems, integrating multiple techniques that cover the path from raw data to high-level interpretation. The concerned applications are art items (sculptures, paintings), biomedical images, satellite images, natural scenes.

- "Audio, Acoustics and Waves" - AAO - conducts activities in signal processing having strong connections with the physical phenomenon that is at the source of the signals, whether acoustical or optical. In digital audio signal processing, the activities span the entire acquisition chain, from capture to analysis or transformation, transmission up to its restitution, with the goal of proposing solutions to the main problems centered around the sound, speech or music, in multimedia applications. In optical information processing, the group contribute to new detection schemes and to the characterization of new materials.

- "Multimedia" - MM- is a group that covers the life cycle of multimedia documents in the framework of a complete chain going from authoring tools for on-line and offline production of multimedia contents to multimodal interaction for the final user; this also includes automated processing like enhancement of degraded pictures, verification of the identity of the user, modification of auditive and visual appearance, image segmentation and pattern recognition. The group also works on techniques that allow the analysis, compression and robust transmission of these media in heterogeneous networks. It also works on the dynamic and distributed adaptation of the transmitted data flow (including meta-data and in particular those concerning the digital rights management) with respect to context, transport conditions and terminal type.

One research topic is common to all groups, this is indexation and data mining. Summarizing and extracting informations from multimodal databases requires statistical tool for learning and mining, which are among the activities of the STA group with a particular focus on text indexation.
and retrieval. Indexing satellite images, extricating informations from primitives to semantic annotations is the main goal of the “Center of Competence”, a joint lab between CNES, DLR and the TII group. This group also develops the same kind of tools for biomedical images and for 3D objects. The AAO group is concerned by many aspects of music information retrieval: identification of rhythms, main melodies, instruments, styles, moods, tonalities either from plain audio or from mixed audio and video. Video signals are also a core activity in the MM group together with complex documents analysis (mixing printed texts, handwritten texts, pictures, graphics) and with multimodal analysis for biometry (voice, faces, fingerprints).

Our most recent recruitments were aimed towards the reinforcement of two topics: the first one is distributed sensor processing; the second one is 3D images and virtual worlds.

<table>
<thead>
<tr>
<th>Permanent staff [Institut ; CNRS ] ; post-docs</th>
<th>[32 ; 11.6] ; 8.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>PhDs</td>
<td>55.6</td>
</tr>
<tr>
<td>Defended PhDs</td>
<td>90</td>
</tr>
<tr>
<td>Defended HDR</td>
<td>8</td>
</tr>
<tr>
<td>Journal papers</td>
<td>291</td>
</tr>
<tr>
<td>Chapters and books</td>
<td>52</td>
</tr>
<tr>
<td>Conference papers</td>
<td>713</td>
</tr>
<tr>
<td>Patents and software</td>
<td>8</td>
</tr>
<tr>
<td>Contractual income 2005–2009 (june) [Private ; Public ; European] (k€)</td>
<td>[5865; 2747 ; 1782]</td>
</tr>
</tbody>
</table>
Chapter 10

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. Vasilescu (CR2-CNRS, 09/05), 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. Ozeryov (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 [IT ; CNRS ] ; post-docs | 10 |
| PhDs | 18 |
| Defended HDR | 2 |
| Journal papers [published, in press] | 53 |
| Chapters and books [published, in press] | 3 |
| Conference papers | 136 |
| Patents and software | 2 |

Contractual income 2005–2009 (june) [Private ; Public ; European] (k€)  560 ; 755 ; 356
10.1 Audio Signal Processing (AudioSig Project)

10 Audio, Acoustical and Optical waves (AAO)

The AAO (Audio, Acoustical and Optical 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)

10.1 Audio Signal Processing (AudioSig Project)

10.1.1 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, . . . ).

10.1.2 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 [1836], efficient algorithms for the adaptive estimation and tracking of the signal subspace components [1828][1831]). An increased effort was also dedicated to sparse signal representations, such as based on Matching Pursuit or Non-negative Matrix factorisation (NMF)[1846], 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 [1859], 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 [1859], use of harmonicity or temporal constraints for music transcription [1833], use of source production or timbre models for source separation [1922]...). 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 [1925],[1983]), rhythmic information tracking (tempo and beat estimation [1825],[1824]), harmonic information estimation (recognition of the chords sequence) and timbre recognition (musical instrument recognition in polyphonic audio [1844]). Source separation approaches were developed for specific music transcription tasks such as drum track transcription and resynthesis [1850] and main melody estimation (by use of a NMF-based source-filter model for separating the singing voice from the musical accompaniment [1923]) but also for specific audio rendering tasks such as stereo signal remastering [1862].

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 [1976], speech emotion recognition [1866],[1840] or TV show structuring) and musical streams (musical instrument recognition [1849],[1857], multimodal audio/video semantic alignment [1849]). 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 [1933]. 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 [1907].

In sound capture, recent work permitted to propose a novel technic for automatic sound field analysis from a network of sensors (microphones) [1944]. 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 [1823]. We produced several variants of this approach, one of them being based on an empirical modal decomposition [1865]. 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 [1822]. 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 [1821].

### 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 [1833]. 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 [2405]. 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 [1842]. 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) [1863]. It was also shown that the high flexibility of the signal representations used in this coder
10. Audio, Acoustical and Optical waves (AAO) 10.1. Audio Signal Processing (AudioSig Project)

allows to tackle various audio indexing tasks (such as beat tracking or musical genre recognition) directly in the transform domain [1864].

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 GMDF (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 [1994], [1835]. 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 [1897]). 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 [1979]; 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 [1961]; 3) we have developed relatively simple models of fricatives that can produce highly intelligible and naturally sounding fricatives in speech synthesis experiment [2019].

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 [1986]. Patent application for each of these two inventions is in progress with help from the CNRS.

10.2 Optical Signal Processing

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

Highlights Institut TELECOM funding on the subject Network 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 [1839], gain media [1860], thin gratings [1872] 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 [1860]. 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 [1872] 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 [1852]. The simple analytical modeling developed for this kind of gratings can be most useful for the design of new devices [1853].
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 [1871]. 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 [2000]. 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.

### 10.3 References

#### 10.3.1 ACL: Articles in ISI-Indexed Journals


10.3.2 ACLN: Articles in Other Refereed Journals


10.3.3 ACTI: Articles in Proceedings of International Conferences


10.3. References

10. Audio, Acoustical and Optical waves (AAO)


10. Audio, Acoustical and Optical waves (AAO)

10.3. References


10.3. References

10. Audio, Acoustical and Optical waves (AAO)


10.3.4 ACTN: Articles in Proceedings of French Conferences


10.3.5 COM: Talks in Conferences Which Do Not Publish Proceedings


10.3.6 OS: Books and Book Chapters

Chapter 11

Multimedia (MM)

Team leader B. Pesquet-Popescu (P).

Faculty M. Cagnazzo (MC, 02/08–), G. Chollet (DR CNRS), C. Concolato (MC, 10/07 –), C. Faure (CR CNRS), J. LeFeuvre (IE), L. Likforman-Sulem (MC), J.-C. Moissinac (MC), C. Pelachaud (DR CNRS, 01/09 –), M. Sigelle (MC), C. Tillier (06/05–09/07)

PhD students H. Bredin (09/04–10/07), S. S. Lin (09/02–06/07), L. Zouari (01/04–04/07), R. El-Hajj (11/05–07/07), T. Hueber (10/06–), B. Pellan (10/06–), B. Elloumi (12/06 –), M. Bendris (10/08–), R. Bayeh (12/03–), C. Angeli (), P. Perrot (01/05–), C. Concolato (10/02–07/07), G. Pau (01/03–06/06), M. Trocan (10/04–10/07), C. Tillier (10/02–06/05), G. Feideropoulou (10/01-04/05), A. Robert (11/04–01/08), C. Bergeron (01/04–01/07), G. Laroche (11/05–05/09), N. Tizon (11/05–), O. Crave (10/05–12/08), I. Daribo (10/06–), T. Magey (10/07–), S. Chebbi (12/06–), C. Yaacoub (04/06–07/09), M. Kaaniche (10/06–), S. Hyniewska (10/08–), Q. Anh Le (07/09 –), J. Huang (10/09 –), C. Greco (09/08 –), R. de Oliveira (01/09 –), M. Kiniaei-Asadi (02/02 – 06/05), B. Rodriguez (02/09–), A. R. Kaced (10/05–10/08), Z. K. Aoul (10/05–10/08)

Post-docs, engineers and sabbaticals R. Landais (05/06–12/07), G. Aversano (01/06–09/06), L. Zouari (04/07–11/08), J. Wei (10/07–09/08), Y-Z. Zhang (03/07–03/08), A. Fraysse (10/06–08/08), S. Brangoulo (01/06–09/06), B. Pellan (–09/06), T. André (10/07–05/08), J. Gautier (06/08–09/09), W. Miled (10/07–09/09), A.M. de Bellaing (10/06–04/07), B. Zalesky (sabbatical 1 month), R. Niewadowski (01/09 – 12/09), A.-M. Pez(10/08–01/10), K. Prepin (01/09–12/09), E. de Sevin (01/09 – 10/10), M. Ochs (09/09 – 08/10), H. Sarria (sabbatical, 01/08–01/09), E. Barney Smith (sabbatical, 2 weeks), E. Sanchez-Soto (05/09–), J. Razik (02/08–08/09), C. Riedinger (11/08–12/09), P. de Cueto (–06/05), A. Amehraye (11/08–)

External collaborators A. Bennazza (SupCom Tunis), R. de Quieroz (Univ. Brasilia), C. Mokbel (UOB, Liban), C. Kermorvant (A2IA), A. Vinciarelli (IDIAP), I. Jermyn (INRIA Sophia), S. Perreau (ITR, Adelaide), E. Bratsolis (Univ. Athens), J. Farah (USEK, Liban), M. van der Schaar (UCLA, USA), G. Plella (Univ. Pompeu Fabra, Barcelona), C. Guillemot (INRIA Rennes)
11.1 Objectives

The research in the “Multimedia” (MM) group covers the life cycle of multimedia documents in the framework of a complete chain, going from authoring tools for on-line and off-line production of multimedia contents to multimodal interaction for the final user; this also includes automated processing like enhancement of degraded pictures, verification of the identity of the user, modification of auditive and visual appearance, image segmentation and pattern recognition. The group also works on techniques that allow the analysis, compression and robust transmission of these media in heterogeneous networks. It also works on the dynamic and distributed adaptation of the transmitted data flow (including meta-data and in particular those concerning the digital rights management) with respect to context, transport conditions and terminal type.

11.2 Main Results

The main research results obtained during the period 2005-2009 are presented below for the research areas of the Multimedia team.

11.2.1 Robust Compression and Transmission of Visual Data

Faculty B. Pesquet-Popescu, C. Tillier, M. Cagnazzo

Main events Best Paper Award IEEE Trans. Circuits and Systems for Video Technology 2006 received by B. Pesquet-Popescu. B. Pesquet-Popescu is a member of the IEEE SPS Multimedia Signal Processing (MMSP) Technical Committee, of the IEEE SPS Image, Video and Multidimensional Signal Processing (IVMSP) TC, Associate Editor for IEEE Trans.on Multimedia, Associate Editor for Elsevier Signal Processing, and was a Co-Chair of the MPEG AHG on Exploration in Wavelet Video Coding (04/05–07/06). She is also a member and Treasurer of the EURASIP AdCom and member of the GdR ISIS administrative committee. In 2006 and 2007 she was also a “rapporteur” for the RIAM program and is, since 2005, an expert for ANR.

Scalable and Adaptive Coding

One of our main topics of interest is scalable video coding [2257], allowing a video system to provide a flexible and robust bitstream, able to be adapted to different transport and visualization conditions. We have studied video coding based on spatio-temporal wavelet decompositions, enabling a natural spatial and temporal scalability. These schemes also have the advantage of easily allowing the implementation of unequal error protection [2056].

In this context, we have been among the first to introduce motion-compensated temporal lifting schemes ([2050, 2029]). New temporal decomposition tools have been proposed like, for example, 3-band temporal schemes [2051], optimized update operators, iterative bidirectional prediction structures, low delay temporal schemes etc. In the spatial domain, we have worked on $M$-band filter banks permitting a fractional scalability [2039], and that can also be adapted to the quantization step and to the content itself. This technique leads to important gains in quality and computational time, and allowed Telecom ParisTech to get a patent and to be in the process of obtaining a second one on the extension of this idea to H.264/AVC and SVC video adaptation. Different other adaptations and optimizations of the AVC/SVC codecs have been implemented in the collaborative project PINGO.

Another application of lifting structures is the construction of content-adaptive filter banks, where the choice between two or several structures is related to a criterion uniquely based on the analysed data. We have proposed [2048, 2057], in collaboration with G. Piella (Univ. Pompeu Fabra), and H. Heijmans (CWI, Amsterdam) an original framework based on semi-norms allowing to provide flexible decision criteria. Recently, we have incorporated rate-distortion criteria in these decompositions. Another approach for designing sparse representations adapted to the video content is based on block-oriented transforms, where we proposed solutions in a standardized framework during the PhD thesis of A. Robert (CIFRE FTRD).

An extension of the previous techniques to multi-view coding and 3DTV compression was performed, in several directions: first, the compression of stereo image pairs and stereo video sequences by joint multiresolution analyses without leakage [2031], and second, the multi-view coding for free-point of view and 3DTV applications [2046]. In this context, we have proposed optimal bitrate allocation in such schemes, as well as pre-, post-processing and compression of depth maps. The disparity estimation in a variational approach with convex constraints is a key point of this theme, which is developed by W. Miled [2038]. It was also extended to dense motion estimation and joint disparity-motion estimation for multi-view coding. Lifting structures for multi-view coding and the application of joint wavelet packets in this framework was done in collaboration with Nuremberg University.

Finally, an important theoretical work, in collaboration with A. Fraysse (now at Univ. Paris-Sud) and J.-C. Pesquet (Univ. Paris-Est), was the study of asymptotic operational rate-distortion curve of Bernouilli-Generalized Gaussian sources, which provide an accurate model for the subbands of different spatio-temporal transforms [2045].

Robust and Joint Source-Channel Coding

In a standardization context of MPEG-4/AVC, we have studied schemes based on “competition” [2034], based on different optimization criteria (PhD thesis of G. Laroche, CIFRE FTRD). Some of the proposed tools have been integrated in the KTA, which is the reference software for a possible future standard (H.265 ?). We have also proposed original solutions for temporal scalability using frame shuffling [2021] (PhD thesis of C. Bergeron, CIFRE with Thalès). In the same collaboration with Thalès, several joint source-channel coding optimizations for H.264/AVC streams have been performed. Video streaming over wireless networks, stream commutation, detection and prevention of congestion, ressource allocation, are the main topics of the work performed by N. Tizon in his CIFRE PhD thesis with SFR [2052].
11.2. Main Results

In the same time, video transmission over error-prone networks may be highly affected by congestion or bottlenecks. A tool allowing to cope with such errors is the multiple description coding (MDC), exploiting the existence of different paths from the sender to the receiver. We developed MDC schemes issued from wavelet frames with reduced redundancy in the spatial and/or temporal domain and proposed several solutions exploiting advanced convex optimization techniques. One of them, called “synthesis frame approach”, allowed us to establish interesting connections with the compressed sensing framework. The MDC paradigm is also considered in collaborative projects like DIVINE (Diffusion de Vidéo et Image vers des termiNaux hEtérogènes, à travers des liens hétérogènes), where we performed unequal error protection for multicast links, or DITEMOI (Diffusion et TElechargement sur lien MOBILE IP), where we work on joint source-channel coding for wireless multi-point to multi-point (Wi-Fi or WiMAX) channels [2041]. In the Sebastian2 project, dedicated to real-time tools for post-production between Paris and San Francisco creation areas, we develop the idea of using MDC for P2P communications and propose new protocols for wired and wireless P2P networks.

Distributed Video Coding

The current development of applications like mobile visiophony raises an increasing interest from the industry for compression techniques with low complexity, and low battery charge, in order to increase the autonomy of mobile terminals. In this context, the distributed source coding paradigm provides original solutions for moving the complexity of video compression from the encoder to the decoder or base station [2024]. Moreover, there is a strong increase of sensor network solutions for videosurveillance, facing similar constraints. In our team, we develop distributed video coding (DVC) schemes, some of them in collaborative projects like ANR ESSOR (codage de SourceS vidéo distRibué), and consider both theoretical and applicative aspects in mono- and multi-view distributed coding, and related multi-terminal concepts. For example, in collaboration with INRIA Rennes, we proposed iterative (turbo-like) decoding of MDC streams with adjacent information [2042]. We have also performed a rate-distortion analysis and error propagation study of mono- and multi-view DVC schemes [2042].

11.2.2 Rich Media, Adaptation and Open Source Software

Faculty C. Concolato, J. LeFeuvre, J.-C. Moissinac

Main events Organization of Distributed Framework for Multimedia Applications 2007, 2nd Best Open Source Software Award in ACM MM 07, Edition of several ISO and W3C standards

Projects ANR MP4MC (01/06–06/07), ANR Radio+(11/08–11/10), Cap Digital PINGO(04/08–04/10), IST FP6 STREP ISIS (09/02–03/05), IST FP6 STREP TIRAMisu (11/03–12/05), IST FP6 STREP DANAE (01/04–06/06), IST FP6 NoE INTERMEDIA (10/06–10/10), ANR Georacing (01/07–03/09), IT JEMTU (01/06–12/08).

Rich Media Representation

The term “Rich Media” [2027] designates the methods, algorithms, tools or technologies required for the processing of the new generation of multimedia content, i.e. content that encompasses natural or synthetic audio-visual material but adds animation and interactive capabilities. Rich Media technologies target a wide range of application domains: from digital TV or radio, to mobile multimedia and the Web 2.0.

The research topic that the team pursued in this area are numerous. Some work has been done towards finding the best representation for such content, as well as designing compression approaches for multimedia scene description languages, with features such as error protection and scalability. Other works focused on improving the visualization of such content [2023], in
11. Multimedia (MM) 11.2. Main Results

particular on mobile phones. Finally, the problems related to the delivery of such content on diverse networks such as broadcast networks have been also investigated. As part of this work, the team is an active participant to standardization bodies such as W3C and ISO. The team has contributed (more than 80 contributions) and participated to the editing of the following international standards: MPEG-4 Systems, MPEG-4 BIFS, MPEG-4 LASeR, and W3C SVG.

Multimedia Adaptation

The adaptation of multimedia content to its context of use (terminal capabilities, network characteristics, user preference) is a very active research topic, with tight link with standardization activities such as MPEG-21 or W3C. The team explores specific problems in the adaptation of multimedia content: adaptation of protected content, adaptation of human-computer interface, and the authoring of adaptable services [2040]. These problems are addressed along different axis either by defining software architectures for such adaptations (in relationship with the ASTRE Team) or by defining methods and languages facilitating the adaptation of multimedia documents.

GPAC Open Source Software

The team maintains an Open Source platform called GPAC [2134], GPAC Project on Advanced Content (http://gpac.sourceforge.net), distributed under an LGPL license. This platform offers various tools for the encoding, the delivery and the playback of multimedia content, ranging from simple audio/video to full-fledge Rich Media. These tools implement state-of-the-art algorithms, methods and protocols from many standards organizations (MPEG, W3C, IETF, ETSI) and are kept in close sync with new industrial deployments. GPAC is used by the academic world (42 citations in journals or international conferences), the industrial world (integrated in several R&D projects) and the internet community (e.g. used for iPod file management). GPAC constitutes the back-bone for the implementation of the team’s work (Rich Media representations and adaptation) and is often demonstrated in conferences or standardization meetings.

11.2.3 Document Imaging and Interaction

Faculty L. Likforman, M. Sigelle, C. Faure

Main events L. Likforman is the General Chair of the Document Recognition and Retrieval (DRR) 2009 and DRR 2010 conferences.


We first showed that the use Dynamic Bayesian Networks for the recognition of either naturally of artificially degraded characters performs significantly better than other current state-of-the-art methods such as Support Vector Machines [2035] (L. Likforman, M. Sigelle). Also, we designed a recognition system for cursive arabic handwriting combining several Hidden Markov Model classifiers using different oriented windows [2028], which proved to be one of the most performant at this moment (L. Likforman, in collaboration with C. Mokbel and R. Al-Hajj, Univ. of Balamand, Lebanon won the ICDAR 2005 competition for Arabic handwritten word recognition). A grant has been provided in 2008 by the A2IA company to the PhD student Anne-Laure Bianne for improving this system. L. Likforman took part in the specification of the TechnoVision RIMES (Written documents recognition and indexing) project goals, as well as in the RIMES evaluation in character and logo recognition (in collaboration with S. Ladjal). Télécom ParisTech was ranked on the first position for logo recognition and second, ex-aequo, for character recognition.

For image restoration with Total Variation based on graph-cuts methods, the PhD of J. Darbon and subsequent work [2025, 2026] have lead to a new methodology for joint restoration of Synthetic Aperture Radar amplitude and phase images for 3D reconstruction of buildings (joint work
11.2. Main Results

with F. Tupin et L. Denis ENSML). A new grant on this subject (funded by DGA/REI) has been accepted and should start soon (in collaboration with J-F, Aujol (CMLA) and J-M. Nicolas). M. Sigelle has also been working in collaboration with W. Pieczinsky (Télécom SudParis), F. Tupin and D. Benboudjema on triplet Markov Random Fields AIMED TO texture analysis and indexing in the framework of the Info@Magic project.

M. Sigelle started a collaboration with I. Jermyn (INRIA ARIANA) and S. Perreau (UNISA Adelaide Australia) on the topics of (discrete) diffusion processes, which can be applied both to modelling of traffic routing in ad hoc networks and to image restoration [2194, 2360]. The studies of C. Faure on documents and images emphasized the role of communication and the visual modality. Digital and digitised documents are processed to facilitate information access. Layout and logical structures are automatically detected in document images or in semi-structured digital documents. Applications were developed for the RNTL project InfRadio for which web documents were adapted to be read and activated on the small screens of mobile devices [2224]. More recently, document image analysis was performed for the digital library medic@ to assist the archivists in indexing and storing historical medical documents. New methods were proposed to structure the images of the pages and to extract relevant components such as the figure and caption pairs [2096, 2094, 2095]. To cope with ancient fonts difficult to recognise by OCR, word spotting methods were proposed to search for word-images similar to query words [2231, 2128, 2129]. These works for medic@ are made in collaboration with the LIPADE (Univ. Paris V). In GEOservice, a joint project between several research teams of the Institut Télécom (C. Faure was prime), the visual modality was involved in a web service. Images were combined with text to provide multimodal egocentric instructions for guiding a mobile user in a building. As a natural complement of the visual modality, the gestural modality was studied in the context of human-computer interaction where the users drew or wrote to communicate [2093, 2223, 2260, 2234, 2086].

11.2.4 Audio-visual Identity/Imposture and Virtual Worlds

Faculty  G. Chollet, C. Pelachaud, M. Sigelle, M. Charbit

Main events  G. Chollet and C. Pelachaud, general co-chairs of IVA’07; C. Pelachaud and T. Boubekeur, co-editor special issue on Facial Modeling, IEEE Computer Graphics and Applications, to appear in 2010; C. Pelachaud co-organizer of a Workshop held in conjunction with AAMAS 2009; she is since 2007 secretary of the Humaine association on emotion; she is part of the selection committee of ANR CONTINT (since 2008), ANR Blanc CSD9 Sciences Humaines et sociales (in 2009).


Two main directions of investigation are present in this theme:

Biometry and Speech/Face Synthesis/Recognition/Verification

The speech group was created in 1983 when Gérard Chollet joined Télécom-ParisTech (called ENST at the time). The focus was centered on coding, synthesis and recognition. In the 1990, speaker verification was added, followed by language identification five years ago. At that time, audio-visual speech and speaker recognition became a topic of interest. The Biosecure network of excellence was an opportunity to promote open-source software for major biometric modalities (face, voice, audio-visual speaker, signature, iris, hand shape...) This led to the publication of
the book ([2269]) and to the development of databases, reference systems and benchmarking protocols ([2849, 2259]). The FP6-Securephone project was an opportunity to integrate audio-visual identity verification on a mobile phone. Audio-visual identification also finds applications in video indexing (InfoM@tic project, PhD theses supported by OrangeLabs,…) Face tracking and super-resolution of faces are issues under study in the ANR-KIVAOU project and are evaluated in the context of the NIST-MBGC campaigns. Speech recognition is still a major problem for our team. It is being experimented in projects such as the ANR-MyBlog3D and the FP7-IP-Companionable in the context of Spoken Dialogue Systems. Initial results on Very Low Bit Rate Speech Coding led to a participation of G. Chollet in start-up companies such as Peer2Phone and Shankaa. Our coder still needs to be improved in terms of speaker and language independence. A similar approach is developed in the context of the ANR OUISPER project aiming at the development of a Silent Speech Interface (driven from tongue and lip movements).

Interaction and Embodied Conversational Agent

We have been developing an interactive platform of an Embodied Conversational Agent GRETA (virtual entity endowed with human-like communication capabilities) (work done within the projects ANR RNTL MyBlog-3D and IP-CALLAS) [2062]. Greta is open source platform under GPL licence (http://www.tsi.enst.fr/~pelachau/Greta/; more than 100 downloads in 1 year; it is used in several international projects as well as material for academic purposes). Two major axes are actually undertaken: the first one relates to nonverbal communicative and emotional behaviour model and the second one focuses on model of the interaction between user(s) and virtual agent(s). Models of communicative and emotional behaviours of ECAs are elaborated within the EU project IP-CALLAS and the national projects ANR CECIL, ANR GV-Lex and ANR IMMEMO. Different aspects of expressive behaviours are being modelled. Our aim is to go beyond the expression model of the six prototypical expressions of emotions that have been mainly considered so far. We are extending our model of expressive behaviours to other modalities than faces such as gesture and gaze [2185, 2135]. Expressions of emotions can correspond to blend of emotions (eg superposition of two emotions) (IP-CALLAS) [2262]. The expression of emotion does not correspond solely to a static facial expression but it corresponds to sequential multimodal behaviours (IP-CALLAS) [2152]; facial behaviours for complex emotions are going to be further defined (ANR CECIL; PhD thesis Jing Huang) from our previous work [2261]; Expressive communicative behaviour for virtual agent and the humanoid robot NAO is being developed within the project ANR GV-Lex (PhD thesis Quoc Anh Le); and finally emotionally-coloured communicative behaviours is being worked out in the project ANR IMMEMO. While in most of our work we based our model on literature and on careful observation of data (IP-CALLAS; PhD thesis Sylwia Hyniewska) [2151], in the project ANR IMMEMO we aim to use learning techniques to motion capture data so as to extract information on the relation between behaviour parameters. Our work on interaction is geared toward elaborating a listener model as well as the emergence of synchronous behaviours between interactants [2256]. Within the STREP SEMAINE we are developing a backchannel model to simulate listener's behaviour in an interaction [2143]. While the project SEMAINE deals with a dyad situation, a user dialoguing with a virtual agent, and is geared toward emotional dialogs, the project NoE SSPNet focuses on social signals. We are elaborating a model of synchrony between interactants of a conversation, synchrony being a sign of engagement. Within SSPNet we are extended our rule-based model to deal with dynamic model. Behaviours of agents are not only specified at a high-level (eg communicative intention and emotion) but they are also dynamically adapted to the user's behaviour [2256].
11.3 References

11.3.1 ACL: Articles in ISI-Indexed Journals


11. Multimedia (MM)  

11.3. References


11.3.2 ACLN: Articles in Other Refereed Journals


11.3.3 ASCL: Articles in Journals Without Review Committee


11.3.4 INV: Invited Talks


11.3.5 ACTI: Articles in Proceedings of International Conferences


T. Petrisor, B. Pesquet-Popescu, and J.-C. Pesquet. Perfect Reconstruction in Reduced Redundancy Wavelet-based Multiple Description Coding of Images. In EUSIPCO ’05, Antalya, Turquie, September 2005.


11.3. References


11. Multimedia (MM) 11.3. References


11.3.6 ACTN: Articles in Proceedings of French Conferences


11.3.7 COM: Talks in Conferences Which Do Not Publish Proceedings


11.3.8 OS: Books and Book Chapters


11.3.9 DO: Journal or Proceedings Edition


11.3.10 AP: Patents, Registered Softwares


11.3. References
11. Multimedia (MM)
11. Multimedia (MM) 11.3. References


Chapter 12
Statistics and Applications (STA)

Team leader F. Roueff (P).

Faculty K. Abed Meraim (MC, on sabbatical leave at University of Sharjah, 06/07–06/09), G. Blanchet (DE), P. Blanchi (MC, 01/09–), O. Cappé (DR CNRS), J-F. Cardoso (DR CNRS), M. Charbit (P, on sabbatical leave at University of Adelaide, 06/07–12/07), S. Clémençon (MC, 10/07–), G. Fort (CR CNRS), A. Garivier (CR CNRS, 10/07–), J. Jakubowicz (MC, 11/08–), E. Moulines (P), C. Lévy-Leduc (CR CNRS), J. Najim (CR CNRS), F. Roueff (P).

PhD students A. Ben Hadj Alaya (10/05–10/08), S. Barembruch (10/07–), T. Ben Jabeur (10/05–), H. Benoudnine (09/06–07/08, phd started at USTO, Algeria), L. Berrice (09/02–04/06), M. Boulé (09/03–09/07), H. Bousbia-Salah (01/05–05/06, phd started at Ecole Polytechnique d’Alger, Algeria), N. Castaneda, (09/04–07/08), J. Cornebise (09/05–06/09, also at Univ. Paris 6), M. Depecker (10/07–), J.F. Germain (09/05–10/08), F. Guilloux (10/05–12/08, also at Univ. Paris 7), Z. Harchaoui (11/05–11/08), M. Karray (09/03–09/07), I. Kacha (01/05–04/07, phd started at Ecole Polytechnique d’Alger, Algeria), M. Kharouf (01/07–), O. Kouamo (09/07–, also at Univ. Yaoundé 1, Cameroon), D. Lahat (09/07–, also at Univ. Tel Aviv, Israel), A. Lung-Yut-Fong (10/08–), N. Mahler (02/08–, also at ENS Cachan), B. Mouhouche (09/02–12/05), N. Sokolovska (11/06–), S. Philipp (10/07–), G. Picard (10/03–12/06), T. Rebaika (10/06–), L. Rigouste (10/03–11/06), W. Souidene (06/03–10/07), T. Trigano (10/02–12/05).

Post-docs, sabbaticals B. Benmammar (postdoc 12 months), P. Etoré (postdoc 8 months, also at CERMICS, Ponts), J. Olsson (Postdoc 10 months), M.S. Taqu (Prof. at Boston Univ., 3 months), L. White (Prof. at Univ. of Adelaide, Australie, 6 months), M. Zetlaoui (postdoc, 1 year), V. Reisen (MC, Vitória Univ., Brazil, 8 months), Samir Attallah (Prof. at NUS, Singapore, 2 months).
12.1 Objectives

The STA team’s main research interest is in the development and analysis of statistical methods for information processing, with applications in signal processing, applied statistics, complex systems and digital communications. The team’s main expertise lies in statistical signal processing and mathematical statistics but also in probability, operation research and, more generally in applied mathematics. The team is also involved in research projects targeting more specific applications, usually in the context of broader collaborations, often supported by funds from the Agence Nationale de la Recherche (ANR). In this context, topics that are relevant to the team expertise include digital communications, astronomical data analysis, security and defense applications (localization, intrusion or anomaly detection), and data mining. In the recent period, the team started to extend its expertise towards statistical machine learning, in particular for ranking and sequential learning applications.

The members of the STA team are actively participating in teaching, typically at the master level and in the fields of probability, statistics, signal processing, machine learning and applied mathematics, at Télécom ParisTech but also in several other Grandes Ecoles of the ParisTech institute (Ecole des Ponts, Ecole Polytechnique, ENSAE) and universities (M2 Modélisation aléatoire at Paris 7 Denis Diderot, M2 Modélisation Vision Apprentissage at ENS Cachan, M2 Ingénierie Mathématique at Paris 11 Orsay, University Paris-Dauphine).

The STA team has developed long term research collaborations with several academic Parisian partners such as Univ. Paris 7 Denis Diderot (LPMA and ADAMIS), Univ. Paris 10 Nanterre (MODALX), Univ. Paris-Est (IGM), Institut d’Astrophysique de Paris, Univ. Paris-Dauphine (Cérémade), research groups in other ParisTech schools (CMBIO, Mines and CERMICS and CERTIS, Ponts) and with the Ecoles Normales Supérieures Ulm (INRIA projects TREC and WILLOW) and Cachan (CMLA). Such collaborations are essential to the team for achieving long term research programs, and, more generally, for exchanging ideas and views within a stimulating academic environment.

These academic relationships parallel industrial partnerships. The latter have been developed in the framework of national research projects (ANR), bilateral contracts, or the funding of PhdDtheses (through CIFRE conventions). Beside favoring our financial autonomy, such partnerships bring practical applications which are helpful for our opening and to remaining active on new research prospects. In the last years, regular industrial partners include the Commissariat à l’Energie Atomique (CEA), Renault, France Télécom R&D and Direction Générale de l’Armement (DGA).

The team enjoys a high national and international recognition with editorial board members in high quality journals such as Bernoulli, ESAIM P&S (E. Moulines) and the Journal of the Royal Statistical Society, Series B (O. Cappé) as well as regular participation as program comity mem-

Finally, members of the team are regularly invited to give talks in national seminars such as the séminaire parisien de statistiques, universities abroad (Hong Kong University of Science and Technology and National University of Singapore, S. Clémençon; probability seminars in University of Bochum and Stanford Univ., J. Najim; seminar of statistics in Cornell Univ. and Université Catholique de Louvain, F. Roueff; seminar of applied probability in Warwick, G. Fort) as well as in workshops or conferences (Isaac Newton Institute, O. Cappé, E. Moulines; 2006 New Developments in MCMC workshop, 2008 Adap'Ski workshop, 2008 SSC-SFDS conference, 2009 workshop on Scaling methods in Warwick, G. Fort; 2009 Physcomnet, J. Najim; 2006 ValueTools workshop, 2006 New Developments in MCMC workshop, 2007 Europarandom Algorithms in Complex Systems workshop, 2008 European Geosciences Union General Assembly, 2008 Sequential Monte Carlo Methods SAMSI workshop, E. Moulines).

12.2 Main Results

12.2.1 Statistical Learning

**Contributors** O. Cappé, A. Garivier, S. Clémençon, C. Lévy-Leduc, E. Moulines, F. Roueff.

**Main events** ANR projects KERNSIG (Learning and kernels for representation and decision in signal processing, 2007–), MGA (Graphical Models and Applications, 2008–), TAMIS (Adaptation, multiple tests, ranking and applications, 2006–2009), BEMOL (Prediction of internet users’ behavior, simulation and collaborative filtering, 2008–); Contracts with France Telecom R&D (two theses) and Renault (two theses).

In the context of the STA team, statistical learning is a new research theme that has been largely developed during the last four years. Our efforts on this aspect have benefited from two recruitments (A. Garivier, S. Clémençon) and from the support of several academic (ANR projects KERNSIG, TAMIS and MGA) and industrial grants. Although recent, the team’s contribution in statistical learning is now recognized, with several team members regularly participating as program comity members to the main conferences of the field (ICML, EMCL, COLT and NIPS). The team also developed strong collaborations on this theme with other teams within the ParisTech alliance (CMBIO, Mines and CERTIS, Ponts) and the INRIA/ENS project WILLOW (F. Bach), with whom we are organizing the popular monthly Paris Tech-Machine Learning reading group, as well as with the CMLA, ENS Cachan group (N. Vayatis).

Since 2006, the team has been active first on kernel methods and more specifically their use for purposes other than supervised classification and, in particular, for signal processing applications (which is the main focus of the KERNSIG project). Our main contributions include a mathematical analysis of kernel-based changepoint detection tests [2559] as well as several extensions to the multiple changepoints and changepoint localization problems.

**Graphical models** is another topic on which the team is active with works on parameter inference for latent variable models used in natural language processing [2472] (in collaboration with F. Yvon, Univ. Paris-Sud 11) as well as online learning algorithms for mixture and hidden Markov models [2393]. The team also worked on several applications of sparse regression and classification using LASSO type procedures [2562] [2468].

**Ranking** has become a very important research theme in the team with a series of works initiated by S. Clémençon in [2399]. The distinctive feature of this approach is to view methods based on the AUC (Area Under Curve) criterion as solving a functional optimization task which
12.2. Main Results

12. Statistics and Applications (STA)

requires adaptive approximation of the optimal ROC (Receiver Operating Characteristic) curve. The computation of confidence bands through resampling for the ROC curve and associated performance criteria has been investigated in [2519, 2539]. Reference [2401] presents the main ranking algorithm, termed tree-rank, and provides a thorough theoretical analysis of its performance. In the context of the ANR project BEMOL, the team also started some preliminary works on related tasks such as collaborative filtering.

More recently, with the arrival of A. Garivier, the team started working on themes related to resource allocation and reinforcement learning, in particular in the context of the PhD thesis of S. Filippi (funded by France Telecom R&D).

12.2.2 Statistical Methods for Astronomy

**Contributors** O. Cappé, J-F. Cardoso, G. Fort.

**Projects** ANR projects COSMOSTAT (Statistical methods for reconstruction and analysis of the cosmic microwave background) and ECOSTAT (Exploration of the cosmic model by statistical methods).

There is a growing interest in statistically and numerically efficient methods for the processing of the complex massive data sets delivered by modern astronomical observatories and surveys. Our team contribute to this domain along several axis, with a strong focus on the Planck space mission of the European Spatial Agency. This mission, which will deliver measurements of the Cosmic Microwave Background (CMB) of unprecedented resolution and sensitivity, requires data processing of outstanding quality. The team has been contributing several challenges in this area.

In the context of the Planck mission, which will deliver multi-frequency sky maps, the first task is to develop component separation methods for extracting the best possible CMB map from these measurements. We contribute a powerful blind separation method which performed very well in the Planck separation challenge [2432] and a fast and robust non-blind method [2404] based on spherical needlets.

CMB maps are spherical maps. Usual space-frequency methods such as wavelet analysis cannot be applied in this context. The ANR project COSMOSTAT has been dedicated to develop multi-scale methods on the sphere. We proposed and studied a promising new tool for this purpose, the needlets, a spherical frame for optimal filtering [2404] and spectral estimation on the sphere [2426, 2415].

Ultimately, CMB data are exploited in a Bayesian framework for the inference of the cosmological parameters (age of the Universe, Hubble constant, etc). The complexity of the models requires a specific approach based on Monte Carlo methods (ANR project ECOSTAT) on which the team enjoys a high level of expertise. In [2458], we have developed an adaptive importance sampling scheme targeted to the specificities of cosmological data.

CMB data enter in cosmological inference via the likelihood of their angular spectrum, which raises several issues in terms of the dependence structure. Again this specificity calls for new methods in statistical analysis on the sphere. We developed exact and approximate likelihoods function for the observation of the CMB sky with missing data [2379] via interpolations methods on the sphere.

The formation of the large scale structures of the Universe by gravitational collapse can be analyzed via the skeleton of the matter density field. Analytical skeleton models require knowing the joint distribution of the field and of all its derivative tensors. This problem is related to the theoretical description of spherical invariants of isotropic fields. We obtained closed form expressions of it via a theory of spherical invariants for isotropic fields [2446].

12.2.3 Statistical Methods for Signal Processing

Projects European REX network NewCom; ANR project MalCom (Random matrices for communications); ANR project SESAME (inference for random matrices and communicaton); Contracts DEMORO (with CS), Blind demodulation (with I2E), Aintercom (with DGA), WAVECOM (one thesis) and France Telecom R&D (one thesis).

Our interest lies in applications of mathematical and statistical tools to performance evaluation and optimization of the physical layer of wireless communications systems. Such approaches have been particularly fruitful in many areas of interest in the last decade.

The first topic of interest is the performance analysis of Multiple Input Multiple Output (MIMO) communications. MIMO systems are widely acknowledged as a mean for increasing the spectral efficiency of wireless communication systems. In order to design efficient MIMO communications, a crucial issue is to evaluate the performance of MIMO transmissions in terms of capacity or outage probability. Random matrix theory is a powerful tool which allows to evaluate such performance indicators [2427, 2428]. Whereas the pioneer works in this field usually assume simplistic communication models, our activity consists in developing new tools for random matrices in order to encompass a wider class of communication models, including realistic propagation channel models and involved transmit/receive architectures.

On the other hand, geo-localization and tracking of base stations and mobile stations of GSM network have been considered (in the context of the DEMORO project, and N. Castaneda’s thesis). This study used both GSM signals with a multiple sensor array and traffic informations and took into account multipath propagation and presence of outliers. Different approaches have been considered: Expectation-Maximization (EM) algorithm and recursive EM for DOA estimation applications but also Monte Carlo methods (or particle filtering) in the context of Bearing Only Tracking [2532, 2531].

A final field of interest for non-cooperative communications is blind signal processing. In this context, it is assumed that the signal coming from an unknown transmitter has been intercepted. The received signal is corrupted by an unknown propagation channel. The aim is to demodulate the received signal in order to recover the transmitted data and to estimate the value of the technical parameters used by the transmitter. In order to achieve attractive performance in terms of Bit Error Rate, our aim is to develop blind demodulation approaches using approximate Maximum Likelihood methods. One of the main stake is to propose methods which are suitable to modulations with high spectral efficiency, that is, in the case where the size of the alphabet used by the transmitter is large (Aintercom project, I2E contract).

12.2.4 Monte Carlo Methods

Contributors O. Cappé, S. Clémençon, G. Fort, E. Moulines.

Projects/Main events ANR project ADAP'MC (Adaptive Monte Carlo Methods); ANR project BigMC (Issues in large scale Monte Carlo)); Organization of the international workshop New directions in Monte Carlo Methods in Fleurance, 2007.

The team has acquired a high reputation in the domain of Monte Carlo methods by working on sequential Monte Carlo methods or particles filtering, Markov chain Monte Carlo methods as well as so-called Population Monte Carlo. Its activity has a strong emphasis on methodological and theoretical developments in Monte Carlo methods.

When applying Sequential Monte Carlo methods (SMC), a well-known problem is the degeneracy of the approximations introduced by the resampling steps. We obtained results on optimal sampling allocation [2455]. We also developed methods for statistical inference in Hidden Markov Models, which exploits the forgetting properties of the conditional hidden chains [2445, 2408, 2412] [coll. with Univ. of Lund, Sweden; and Univ. of Jerusalem, Israël].

The efficiency of the Markov chain Monte Carlo (MCMC) methods relies on the tuning of design parameters. New algorithms are based on self-tuning of the parameters on the fly without relying on a priori expert parameter tuning, thus yielding to adaptive MCMC algorithms. We
developed techniques to identify the optimal values of these design parameters \[2420\]. We obtained results on the asymptotic behavior of these adaptive procedures \[2370\] [coll. with Univ. of Illinois, US; and Univ. of Bristol, UK].

**Population Monte Carlo** methods are designed as generic self-adaptive importance sampling algorithms. The goal is thus to calibrate the best fitting proposal. We developed an adaptive method for an automatic computation of the optimal proposal among a class of parameterized importance functions \[2391\].

Developing proper theoretical tools is an important issue for Monte Carlo methods: studying the simulation problems by using theoretical tools used in the theory of Markov chains and particle approximations allows to identify the key convergence bottlenecks and to propose the appropriate methodological approaches to solve them. We obtained results in the Markov chain theory \[2409, 2407, 2630, 2413\], in limit theorems for weighted samples \[2411\] and in output analysis for Markov models \[2397\] using bootstrap methods.

### 12.2.5 Time Series

**Contributors** M. Charbit, S. Clémençon, C. Lévy-Leduc, E. Moulines, F. Roueff.

**Projects** ANR projects OSCAR (Overlay network security: characterization, analysis and recovery) and SARAH (Standardization of high-definition audio remastering); Contracts with CEA (one thesis) and CSA; Participation to the European IP project SECOQC.

Statistical inference for time series and, more generally, for stochastic processes is a wide area. The research activities of the team in this domain covers long standing problems in statistical signal processing and new directions in spatial statistics. These topics are often motivated by applications that are also of interest to the team.

A first topic is concerned with **time frequency analysis** of time series based on a study of the asymptotic statistics in a semi-parametric or non-parametric framework. Our expertise in **long range dependence** has been increased, in particular by a thorough analysis of semi-parametric Wavelet methods \[2442, 2443, 2450, 2449\], coll. Boston Univ.). Specific domains of application have been considered, such as financial time-series \[2414, 2416\] and teletraffic data \[2417, 2470\]. Other subjects in time frequency analysis have been considered such as frequency estimation for irregularly sampled series \[2434\] in a non-parametric framework and missing-value estimation for an AR process applied to DNA microarray data \[2396\], coll. Univ. of Sydney).

A second topic of interest for the team is **change detection** by statistical methods and their applications. We have been working on anomaly detection in Internet teletraffic data (ANR-RNRT project OSCAR, \[2435\]) based on non-parametric statistical methods. An online algorithm \[2632\] has been proposed and implemented in a platform dedicated to anomaly detection in the Internet. On the methodological side, we proposed new change detection methods based on LASSO for automatically selecting the number of changes and kernel methods for change detection using unspecified features \[2562, 2559\].

Our activities include theoretical studies of specific stochastic processes arising in applied probability and/or having a strong impact on specific applications. We have been interested in **spatial point processes** for modelling natural images using geometrical models (coll. with TII team \[2673, 2651\] and quantum key distribution networks \[2457\], SECOQC project, coll. with MIC2 team). With a particular emphasis on the time evolution of spatial point processes, we also considered stochastic epidemic models \[2462, 2381\]. The **pileup models** appear naturally in several measurements context such as spectrometry and fluorescence. We proposed statistical methods which take into account the pileup phenomenon rather than avoiding it leading to new algorithms for processing such measurements \[2454, 2633\]. A coll. with Univ. of Lille and Michigan State Univ. yielded new results on the path properties of \(\alpha\)-stable fields \[2374, 2459\]. We also studied the **extremes** (tail properties) of Markov chains \[2380, 2382\], which are of interest in risk management.
12.3 References

12.3.1 ACL: Articles in ISI-Indexed Journals


228
12.3. References


12.3. References


12.3.2 ACLN: Articles in Other Refereed Journals


12.3.3 INV: Invited Talks


12.3.4 ACTI: Articles in Proceedings of International Conferences


12.3. References


12. Statistics and Applications (STA) 12.3. References


12.3. References

Allemagne, June 2009.


12.3.5 ACTN: Articles in Proceedings of French Conferences


12.3.6 COM: Talks in Conferences Which Do Not Publish Proceedings


12.3. References

12.3.7 OS: Books and Book Chapters


12.3.8 AP: Patents, Registered Softwares

Chapter 13

Image Processing and Interpretation (TII)

The Image Processing and Interpretation Group (TII) of the TSI department includes research projects dealing with images and 3D objects, and the Center of Competences in information extraction and image understanding for earth observation (CoC).

Team leader  F. Schmitt (P) until October 2008, then I. Bloch (P).


Supporting permanent staff  (shared with other groups): D. Asselineau, S.C. Barrière, B. Nabati.

PhD students Defended: D. Cherifi (03/05), S. Ladjal (03/05), T. Tung (06/05), F. Duguet (06/05), S. Homayouni (12/05), P. Soler (03/06), D. Girardeau-Montaut (05/06), F. Bretar (06/06), F. Rossant (10/06), C. Valade (12/06), F. Cellier (01/07), G. Peters (06/07), A. Moreno (09/07), J.-F. Goudou (10/07), L. Gueguen (10/07), J. Gerhardt (10/07), C. B. Akul (11/07), B. Zhang (11/07), L. Bin (12/07), A. Bhattacharya (12/07), C. Millet (01/08), H. Khotanlou (02/08), T. Hurtut (03/08), I. Kyrgyoz (05/08), J. Delliere (06/08), N. Bonnier (09/08), M. Costache (09/08), A. Kermi (10/08), P. Lopez Quiroz (11/08), A. Bailland (12/08), E. Belbarbi (01/00), X. Perretton (01/09), A. Ghaleb (02/09), M. Liéon (03/09), O. Nempont (03/09), J.-B. Bordes (04/09), H. Chaabouni (06/09).

Current: E. Aldea (10/06), C. Angeli (01/07), J. Anquez (02/06), S. Audièr (10/08), J. Baussé (12/06), P. Birjandi (01/08), H. Bizot (10/08), P. Blanchart (10/08), E. Bughin (10/07), J. Caron (10/08), J. Chen (doctoral stay, 1 year), N. Chenuard (10/06), D. Craciun (10/06), C. Deledalle (10/08), V. Duval (09/08), G. Fouquier (10/08), G. Ferraioli (doctoral stay, 1 year), B. Galerne (10/08), I. Ghoshal (11/08), D. Hadidi (11/08), G. Hochard (11/07), C. Lemen (10/05), G. Lebreu (10/06), D. Lesage (10/05), C. Mallet (09/07), M. Marim (12/07), D. Martinez (10/06), F. Mosca (10/07), T. Napoléon (10/06), G. Palma (02/07), B. Petipas (10/08), J. Rabin (10/06), S. Redko (05/06), A. Shabou (10/07), A. Simac (10/06),

\[1\text{Francis Schmitt received his engineering diploma from the Ecole Centrale in Lyon, France, in 1973 and in 1979 was awarded a PhD from the University of Paris VI (Pierre et Marie Curie). From 1973 up to his sudden death in October 2008 he was a member of the faculty at Télécom ParisTech (Ecole Nationale Supérieure des Télécommunications, last holding the rank of full professor in the Image and Signal Processing Department in which he headed the Image Processing Group. His main interests were in computer vision, 3D modeling, image and 3D object indexing, computational geometry, multispectral imagery and colorimetry. He authored or co-authored nearly 150 publications in these fields.} \]
13.1 Objectives

The objective of the group is to develop methodologies and theoretical tools for image, scene and 3D object processing and interpretation. The main approach consists in solving globally complex problems, based on rigorous theoretical bases, and integrating multiple and complementary techniques, for deriving interpretations from data. Applications focus on medical imaging, aerial and satellite imaging, natural image analysis. Contributions of the group can therefore be found at the theoretical level (knowledge and information representation and modeling, at various levels and in 2D as well as 3D, processing, interpretation and reasoning on spatial data), at algorithmical level (in particular to implement the developed models for large and complex data sets), and at applicative level. The group is now well recognized, in both academic, institutional and industrial domains. It has numerous collaborations with other universities, and is supported by grants and contracts. The different research activities are closely linked together, which is one of the strong features of the group.

Over the last four years, the team has benefited from the appointment of three CNRS researchers and one associate professor, strengthening research axes in indexing and mathematics for image processing and computer vision, and in computer graphics. The good reputation of the group and its visibility, in France as well as at international level, are confirmed by the number of publications, but also by the number of collaborations, mentioned below for each research axis, and by its attractiveness for CNRS candidates, post-docs and PhDs.

The scientific animation of the team includes a general seminar and several specific ones (medical imaging, compressed sensing, radar imaging, CoC seminar...). PhD candidates are invited to present their work at the end of the first year of their PhD, so as to gather comments from the whole team and initiate discussions among them, thus favoring cross-fertilization of ideas. We also pay attention to the accompanying process of the PhD theses, beside the direct scientific supervision, including a help to prepare their future.

Faculty [EC, CNRS] | [12 ; 2,5]
---|---
PhD students | 27,4
Post-docs, engineers and sabbaticals | 3,5
Defended PhD theses | 37
Defended HDR | 1
Journal papers [published, in press] | [75, 11]
Papers in conference proceedings | 302
Chapters and books | 26
Patents and software | 1
Grants [public, private, european] (k€) | [2705, 662, 97]
The team is also strongly involved in teaching, both at undergraduate level and master level, at Télécom ParisTech and in partner universities. It is responsible for several master programs in image processing and its applications to medical imaging and satellite imaging, thus ensuring a strong link between teaching and research.

13.2 Main Results

The main research results obtained during the period 2006-2009 are presented below for the research areas of the TII team, both from a theoretical and methodological point of view and from an application perspective.

13.2.1 Knowledge Representation and Spatial Reasoning

Faculty I. Bloch, M. Campedel, H. Maître.

Main events RFIA 2008 (program chair) and edition of a special issue of the I3 journal, with selected papers.

Projects Collaborations: J. Atif (Univ. Antilles-Guyane), R. Cesar (Univ Sao Paulo, Brazil), C. Hudelot (ECP), J. Inglada (CNES), J. Lang (IRIT and LAMSADE), N. Milisavljevic (RMA, Brussels), R. Pino-Perez and C. Uzcategui (Univ Los Andes, Merida, Venezuela), F. Rossant (ISEP), L. Laborelli (INA), S. Dubuisson (LIP6).

Spatial reasoning in images requires to develop tools for representing spatial information, both for objects and their spatial relations, and for reasoning on this type of information. Uncertainty and imprecision management, as well as fusion of heterogeneous information are central in our work. As the continuation of previous work, we proposed models for representing spatial relations based on fuzzy sets theory \[2643, 2648, 2646\]. Recently, we proposed new definitions of fuzzy connectivity, based on the notion of hyper-connectivity, and dealing properly with the fuzzy sets semantics and with continuity issues \[2692\]. The associated algorithms are based on tree representations, that make filtering and other processing tractable. We also addressed the modeling of complex relations such as “parallel” and “across”, again using fuzzy mathematical morphology. A new orientation of our work deals with the modeling of bipolar spatial information, in order to represent both positive and negative information. We proposed a novel approach, based on mathematical morphology on the complete lattice of bipolar fuzzy sets to represent and manipulate such information \[3035\].

Besides knowledge representation aspects, we addressed the reasoning component of spatial reasoning from different points of view. We developed an ontology of spatial relations, which was used to enrich a part of the FMA\[2\](medical ontology) concerning brain structures. Fuzzy models of spatial relations define the semantics of ontology concepts and their representations in the spatial domain contribute to reduce the semantic gap. This provides a promising way for using the enriched ontology to guide the recognition of image structures \[2677\]. For each particular application, the semantics of the spatial relations (in particular the shape and parameters of the membership functions) are learned on a database of examples. Other work on ontologies, in the domain of satellite imaging, are carried out in DAFOE project (see Section 13.2.5).

These models have also been integrated in graphs representing image structures (objects and spatial relations between them). Reasoning schemes in these graphs have been designed, in order to find optimal paths providing an ordered sequence of objects to be recognized, each object being processed based on the previously processed objects in the sequence and on spatial relations with respect to them. The optimality is defined in terms of spatial relations and saliency computed from the actual data \[2848\]. As a novel contribution, we also integrated these models in constraint networks, and expressed the recognition process as a constraint satisfaction problem,

\[\text{http://sig.biostr.washington.edu/projects/fm/}^2\]
for which we derived specific propagators for each spatial relation in order to reduce the domains of the solutions [2921]. Finally, fuzzy spatial relations are integrated in particle filters for tracking objects in video sequences (collaboration with the LIP6). This new contribution shows a better behavior than classical particle filters in case of abrupt changes in the trajectory.

Our work on fuzzy mathematical morphology has led to the development of new transformations, for defining fuzzy influence zones and skeleton by influence zones, with applications to interpolation between fuzzy sets [2645]. These transformations have also been developed in a logical framework (in collaboration with R. Pino-Perez, C. Uzcategui and J. Lang), with applications to mediation and negotiations [2771].

Finally, our work on information fusion deals with fusion of spatial relations, fusion of defect detectors for digital film restoration (with INA) [2700] and fusion of fuzzy musical rules, which led to higher recognition rates in various musical scores than commercial softwares (with ISEP) [2699]. We also have a long collaboration with the RMA in Brussels for fusion in the domain of anti-personnel mine detection for which we proposed original methods based on belief functions and possibility theory [2686].

13.2.2 Machine Learning and Image Retrieval

**Faculty** The whole group is involved in this research axis.

**Projects** Infom@gic in pôle CapDigital, ANR 2006 AVEIR and DAFOE projects, European project K-Space. Collaborationss with J.-Y. Audibert and R. Keriven (Ponts ParisTech).

In 2006-2007, a new research direction, spreading across various themes in the TSI department, has emerged. In the TII group, it concerns indexing of multimedia documents. By indexing, we mean the analysis of images or documents contents, in order to facilitate their massive exploration. Indexing is strongly linked to the mining operation an end-user may need. Research in this domain benefits from methodological advanced developments (modeling, adaptive learning depending on the type of images), in strong connection with STA team, and from a better knowledge and exploitation (fuzzy reasoning, visual or domain ontologies). Indexing of 3D models was also studied, based on either 2D views or purely 3D information, using kernel approaches for estimating joint density distributions [2635, 2636], and using Reeb graphs [2703].

For 3D object recognition in biological vision, we found that view-specific and 3D-model based representations are used by human observers [3046, 2695]. Finally, mining strategies for large image databases are developed, based on relevance feedback.

Spatial relations have been exploited in this context for recognizing regions of an image and providing a linguistic description of its content (with CEA-LIST). Classification and image mining are also addressed using marginalized graph kernels, and have contributed to the Infom@gic project.

A software platform, PLATO, is being developed with the aim of organizing, centralizing and handling multimedia data (images, sounds, videos, but also processing tools and processing results), in collaboration with AAO team.

The goal of the UrbanView project (partners LIP6, EADS, THALES, etc.) is to design machine learning techniques for multi-camera object (car, person,...) tracking, retrieval and traffic surveillance. Two different scenarios were considered, synchronous and asynchronous, depending on the fact that objects and tracks are matched using overlapping or non overlapping cameras. In this work, we introduced a framework for multi-view object matching and tracking based on kernel canonical correlation analysis. Our method is purely statistical and encodes intrinsic object appearances while being view-point invariant.

Further collaborations, mainly with Ponts ParisTech, include kernel design for object-based image retrieval. The goal is to incorporate many properties (invariance, context, etc.) in order to achieve object matching and retrieval. Theoretical properties, about the positive definiteness of these kernels and their convergence to a fixed point, were proved together with experiment validation on widely used databases including Corel and Flickr [2947].
3D retrieval has recently emerged as an important boost for 2D search techniques, by its several complementary aspects, for instance, enriching views in 2D image datasets, overcoming occlusion and serving in many real world applications such as photography, art, archeology and geo-localization. In this work, we introduced a complete “2D photography to 3D object” retrieval framework which, given a (collection of) picture(s) or sketch(es) of the same scene or object, allows us to retrieve the underlying similar objects in a database of 3D models. The contributions of the method include (i) a generative approach for alignment which is able to find canonical views consistently through scenes/objects and (ii) the application of an efficient but effective matching method used for ranking. The results are reported through the SHREC benchmarking consortium and evaluated/compared by a third-party, showing clearly the good performance of the proposed framework with respect to the other participants [2917].

The AVEIR ANR project is about combined text and image retrieval joining LIP6, LSIS and LIG; its goal is to design machine learning techniques in order to learn the relationships between text and images and perform inference (i.e., image annotation). The members of the consortium are actively collaborating and participating in different evaluations and challenges including ImageClef 2008 and 2009; they submitted a common run ranked 2nd among 25 international experienced teams working on the same topic.

Another research topic is to use manifold learning techniques (graph Laplacian and diffusion maps) for relevance feedback based image retrieval. A new graph Laplacian technique was introduced which makes it possible to robustly learn the embedding of a manifold enclosing an image database, via diffusion map [2949, 2944]. The approach is three folds, it allows us (i) to integrate all the unlabeled images in the decision process (ii) to robustly capture the topology of the image set and (iii) to perform the search process inside the manifold. This technique shows a clear and a consistent gain with respect to state of the art relevance feedback approaches on standard databases. The graph Laplacian technique was also used for dimensionality reduction and applied to large scale image database “visualization”.

Finally, we recently addressed the problem of image queries in large databases from user sketches (binary strokes). We proposed a new descriptor [2838] for fast large scale search and integrated the so-defined search engine within a variational image compositing tool [2839].

In this part of our activities, a core feature concerns kernel-based statistical methods which allow taking into account invariance and contextual properties for object matching and recognition in images and video sequences. The main goal is to integrate additional information about geometry, textual relations and invariance properties in the kernel definition. Theoretical properties of kernels have then to be proved in order to use them for machine learning and dimensionality reduction. Taking into account the transductive aspect is important, via the introduction of prior information in a weakly supervised manner and will lead to increased performances in recognition and interpretation tasks. Multiple applications can be anticipated, such as scene recognition, interactive search and navigation in multimedia generic and specific databases, within ongoing projects such as ANR AVEIR.

13.2.3 2D and 3D Mathematical Modeling

Faculty A. Almansa, T. Boubekeur, J. Delon, Y. Gousseau, S. Ladjal, H. Maître, F. Roueff, F. Schmitt.

Projects European project MUSCLE, ANR Otarie, ANR Freedom, ANR CeCil, ANR NatSim. Collaborations with L. Alvarez (U. Gran Canaria, Spain), J.-F. Aujol (ENS-Cachan), J.-M. Morel (ENS-Cachan), L. Vese (UCLA), V. Caselles (UPF, Barcelona), S. Durand (U. Paris Descartes), M. Lindenbaum (Technion, Israël), P. Musé (U. de la República, Uruguay), A. Sobolevskii (Poncelet Lab., Moscou), T. Buades and A. Desolneux (U. Paris Descartes), S. Masnou (Paris 6), Mila Nikolova (ENS-Cachan), I. Lyuboshenko (PhaseView), M. Alexa (CG Lab, TU Berlin).

Main events International Color Consortium (ICC), digital printing days (March 2009).
Texture and Natural Images Modeling  This research theme deals with the stochastic modeling of natural images. First, generative models taking into account scaling phenomena in natural images have been developed. These models (dead leaves, shot-noise, transparent models) are grounded in the theory of marked point processes, whose marks are geometrical structures [2651]. In particular, we have shown that some models enable the simultaneous representation of geometry and scaling properties in natural images [2673]. More recently, we applied such models to image and texture synthesis. A second research direction is concerned with the mathematical analysis of variational methods for image restoration, and in particular the influence of such methods on the geometry and textures within images. In particular, we have shown that the popular TV-L1 model is equivalent to some morphological filtering [3061]. Another contribution is concerned with the variational decomposition of color images.

Mathematical Methods for Image Analysis and Computer Vision  These last years, we have developed or applied various mathematical tools for the analysis indexing or matching of images. Among these tools, let us first mention optimal transportation equations. These equations enable the definition of metrics between weighted features and yield elegant ways to compare images. Another methodological aspect of our researches deals with a contrario methods, developed by Desolneux et al. to automatically fix detection thresholds for image analysis. In particular, we applied these methods to the problem of image matching. Among the other tools that we have investigated and applied, let us mention topographic maps, scale spaces, and deformable models.

We first proposed solutions to the decision problem for shape matching [2690]. We also have developed a complete chain for the matching of images from local descriptors (such as SIFTs). This procedure encompasses the descriptors themselves, a transportation metric adapted to circular histograms to compare them, an unsupervised matching criterion and a validation, RANSAC-like step [3070][2936][2937]. Another research direction concerns the indexing of satellite images, invariant to resolution changes [2685][2684] or relying on morphological tools [2683]. More recently, we have proposed an original method for the indexing of texture, respecting a wide range of geometrical and radiometric changes [3073]. This method can be seen as an extension of the classical granulometry from mathematical morphology. We also took interest in the indexing and matching of museal artworks, first through the unsupervised comparison of the color composition of images [2678], and then through the automatic analysis of artistic hand drawings [2873]. In the domain of artwork, original contributions on multispectral imaging have been developed for high quality image acquisition [2696]. A mathematical framework for spatial and color gamut mapping has also been proposed, leading to adpative algorithms with real applications for color printing [2898][2775][3058]. In the domain of aerial image matching, we have shown under which conditions a matching is licit, with a precision of a tenth of a pixel. This enables one to develop stereoscopic vision systems with very small b/h [3042][2662][2941].

Restoration of Old Movies  As part of a research project (FREEDOM JCJC ANR project), we have proposed several contributions in the field of movie and videos restoration, in collaboration with researchers from the CMLA (ENS Cachan) and J-L Lions Lab (Paris 6 University). In these contributions, various tools have been used (statistical tests, variationnal approaches, copy-paste methods, patch-based methods, Fourier analysis) and both theoretical and applied points have been tackled, as for instance: the automatic combination of patch-based methods and geometrical interpolation for image inpainting [3050]; the variational interpretation of copy-paste methods [3055]; the automatic detection of occulting defects (dust, scratches) and the restoration of local radiometric problems [2659], for which it has been shown that a precise motion estimation was not necessary, etc. Some of the algorithms developed by the team should soon be made available as plugins for standard movie processing softwares.

At the same time, we also took interest in superresolution and irregular sampling problems. A first direction deals with subspace methods. We continue researches previously developed several years ago at the TSI department and include regularity constraints to circumvent the intricate
problem of source separation in the image superresolution context [2851]. A second direction uses total variation for restoration and superresolution in the case of irregular sampling [2667].

Finally, our work on phase reconstruction for optical waves was pursued, for incoherent cameras and several axial views (in collaboration with PhaseView and I. Lyuboshenko).

**3D Computer Graphics** The group has a long history in 3D image and object acquisition, modeling, processing and interpretation. A new focus of our research activities concerns computer graphics, with the arrival in fall 2008 of Tamy Boubekeur. We have mainly focused on efficient and scalable methods for geometric modeling and realtime rendering.

We have introduced TopStoc [2652], a fast mesh simplification algorithm. The two main components are stochastic vertex selection and re-indexing of triangles. The probability for vertex selection depends on a local feature estimator, which prefers areas of high curvatures but still ensures sufficient sampling in flat parts. Re-indexing the triangles is done by breadth-first traversal starting from the selected vertices and then identifying triangles incident upon three regions. Both steps are linear in the number of triangles, require minimal data, and are very fast, while still preserving geometrical and topological features. Additional optional processing steps improve sampling properties and/or guarantee homotopy equivalence with the input. These properties provide an alternative to vertex clustering especially for CAD/CAM models in the areas of previewing or network graphics.

Ambient occlusion captures a subset of global illumination effects, by computing for each point of the surface the amount of incoming light from all directions and considering potential occlusion by neighboring geometry. We have introduced an approach to ambient occlusion [2938] combining object and image space techniques in a deferred shading context. It is composed of three key steps: an on-the-fly voxelization of the scene, an occlusion sampling based on this voxelization and a bilateral filtering of this sampling in screen space. The result are smoothly varying ambient terms in occluded areas at interactive frame rates without any pre-computation. In particular, all computations are performed dynamically on the GPU while eliminating the problem of screen-space methods, namely ignoring geometry that is not rasterized into the Z-buffer.

As for perspectives, scalable geometric optimizations such as our simplification algorithm, offer a nice framework for large scale problems involved in Cloud Visual Computing. We will study deeper how such methods can be used in a realtime context, to allow high resolution dynamic geometries with all-frequencies content in interactive applications. On the other side of the computer graphics pipeline, we have built our ambient occlusion algorithm on an hybrid object-image space basis. This opens a way toward a more general hybrid rendering engine capable of achieving complex illumination effects such as color bleeding, subsurface scattering or even global indirect reflections, while using volumetric representations as a medium between object and image spaces. This also emphasizes the current convergence between computer graphics (object space) and computer vision (image space) methods.

**13.2.4 Medical Imaging**

**Staff** E. Angelini, I. Bloch, T. Boubekeur, J. Delon.

**Main events** IEEE ISBI 2008 (program chair, finance chair, organization).

**Projects** collaborations within ANR MARIO, INCA, Fondation Santé et Radiofréquences FEMONUM, GET, MINIARA (pôle de compétitivité MEDICEN) projects, J. Rolland and A. Santanam (Univ. South Florida), A. Laine (Univ. Columbia, NY), J. Darbon (UCLA), Y. Petegnief, D. Hasboun and H. Duffau (CHU Pitié-Salpêtrière), IFR 49, E. Mandonnet (CHU Lariboisière), B. Devaux (Ste Anne hospital), C. Adamsbaum (Hôpital Saint Vincent de Paul), E. Moussieux (HEGP), M. Paques and S. Tick (XV-XX Hospital), C. Prunier (CHU Tours), A. Herment and F. Frouin (INSERM, LIF), A. Osorio (LIMSI), M. Teichmann (INSERM), P. Moireau and D. Chapelle (INRIA/MACS), Jean-Christophe Olivo-Marin (Institut Pasteur), F. Rossant (ISEP), O. Gérard (Philips, GE), S. Muller (General Electric), J.F. Stevenet and S.
13.2. Main Results

Hammer (Segami), J. Wiart (FT R&D), Gareth Funka-Lea (Siemens), H. Kafrouni, C. Diaz, and A. Guimond (Dosisoft), R. Ferrand (CPO), Volcano, Fovéa, V. Miette et L. Sandrin (Echosens).

In order to address difficult problems in medical imaging related to the huge size of the data, the complexity of knowledge and information to be processed, the inter-individual variability and the potential presence of pathology, we develop approaches in which knowledge representation plays a central role. Our research focuses mainly on segmentation, recognition and longitudinal analysis of pathological images, in particular for oncology and tumoral pathologies. While the analysis of normal images for several years has led to a very good understanding of the image content in several imaging modalities, the extension to pathological case is difficult and methods relying only on shape and appearance often reach their limits. Our proposal for modeling anatomical knowledge is to make an intensive use of spatial relations (see Section 13.2.1), formalized using fuzzy mathematical morphology, ontologies and graph-based representations. Their integration in deformable models and the analysis of their stability among individuals and in case of pathologies led to robust and accurate segmentation and recognition results [2657, 2689].

Detection of tumors in MRI data has been addressed using a combination of fuzzy methods and deformable models, and was evaluated on a large data base [2679]. Recognition of the normal structures could then be addressed using the same method as for normal cases, since most spatial relations remain stable in pathological cases. We introduced more flexibility in the spatial constraints, for the relations that are prone to strong changes due to the presence of the tumors [2748]. Another approach for the segmentation of multi-modal images has been proposed, based on an extension of the multi-phase level sets model to the multi-channel case. As for the longitudinal follow-up of tumors, a new method for normalizing MRI images and a statistical analysis of difference maps have been developed, which allow designing a framework for automatic quantification of tumoral growth.

All these results have been incorporated in a graph representing both generic knowledge and information extracted from images, with the aim of enriching digital patient records [2694]. The graph representation is also exploited in a web application dedicated to medical teaching, developed in collaboration with D. Hasboun.

Our work on pathologies does not only focus on brain imaging. In thoracic oncology, we have improved our previous non-linear registration methods with a new formalism for constraining the deformations of the pathologies during the registration, while preserving a continuous deformation field (project with Segami) [2653]. Moreover, a breathing model developed at the University of Central Florida was integrated in the registration, thus guaranteeing physiological consistent deformations [2688]. A new project was initiated with Dosisoft (within the "Pôle de compétitivité" MEDICEN) on the segmentation of CT and PET images for radiotherapy applications.

In mammography, our collaboration with General Electric has led to one of the first methods for micro-calcification and mass detection on data obtained with new 3D digital mammography techniques. Recently, new methods for denoising such images and for detecting convergence areas using an a contrario method have been proposed.

In the context of a collaboration with Columbia University (New York, USA), several projects were carried out focusing on the processing of 3D real-time ultrasound data for the characterization of cardiac function (one NIH project, collaborations with Philips Healthcare and Siemens). Dedicated speckle-tracking algorithms and real-time deformable models formulated with active graph functions [2833], in prolate spheroidal coordinates, have lead to novel methods for extraction of myocardial surfaces and tracking of myocardial points. Extensive clinical studies on dog experiments [2836] have been performed to precisely assess the accuracy of local myocardial deformation quantification on ultrasound data. In addition, we also have an on-going collaboration with INSERM LIF group, focusing on the segmentation and quantification of cine and delayed-enhancement MR images, leading to quantitative results on myocardial infarct transmurality and on the estimation of regional mean transition times and radial velocities [2666]. In vascular imaging, a collaboration with Siemens Corporate Research (Princeton, USA) led to the development of several novel methods for the tracking and segmentation of coronaries in high.
resolution CT images, using morphological image filtering and tracking with minimal paths and particle filters. Very accurate results have been obtained on an publicly available data base. A collaborative project with Columbia University and Volcano had focused on the exploitation of multiscale texture-based brushlet analysis for the decomposition of intra-vascular ultrasound (IVUS) data and the extraction of coronary arteries lumen borders.

Regarding the modeling of the human body, which concerned mainly adult and children head until now [2707], a new direction was taken and focuses on fetus modeling, based on MRI and US data (in collaboration with Saint Vincent de Paul Hospital and France Telecom R&D). A variational segmentation method has been developed for 3D US data, taking into account the statistical distributions of maternal and fetal tissues. In MRI, the segmentation is based on anatomical knowledge, driving a graph-cut segmentation. Meshed models are then derived from the segmentation using recent geometry processing methods derived from mesh-based computer graphics techniques and embedded in a synthetic woman body. Preliminary results on dosimetry simulations show that the local and the whole body specific absorption rates are lower in the fetus than in the mother and that they depend on position and morphology but not on gestational age. A common lab with France Telecom R&D (Orange Labs) is currently being launched on this topic.

Recently, a collaboration with ISEP and the XV-XX Hospital was initiated on the analysis of OCT and adaptive optics images of the retina, which led to the development of an automated method for segmenting all layers of the retina. The proposed method was a basis for a preliminary quantitative study of variation of the morphology of foveal and perifoveal layers within a population of healthy subjects.

Finally, a new research axis has been initiated recently, in collaboration with Institut Pasteur, on biological imaging, for tracking and compressed sensing applications. Original multiple hypotheses tracking methods have been proposed, by joint estimation of kinetic and image models [2809, 2806], and CS-based denoising and acquisition protocols have been designed for improved image quality with reduced acquisition times, in the context of fluorescence imaging [2903].

13.2.5 CoC

Faculty M. Campedel, M. Datcu, H. Maître, S. Rital, M. Roux, T. Tanzi.


The joint CNES-DLR-Télécom ParisTech competence center (CoC) was created in June 2005. Its research activities focus on information extraction and image content understanding, for both satellite and optical images (Télécom ParisTech), and SAR images (DLR). It regularly involves about five permanent researchers and 10 PhD candidates. Both theoretical [2684, 2685, 2674] and applied researches are carried out and deal with image indexing and their usage. The images are characterized by their large size, with an important semantical variety of scenes, and their huge number (the Pleiade satellites will send 450 images per day with a 70cm/pixel resolution in 2010!). It becomes urgent to develop (semi)-automatic methods for rapidly accessing the contents of these images. Our current research directions allow us to describe the image content in terms of colorimetry, geometry, texture, and semantics, by using learning methods or pattern detection from which semantical objects are derived (river or road networks, buildings). The learning can be performed either interactively and adapted to the user (photo-interpreter) using relevance feedback, or using statistical inference methods. Finally, the CoC is involved in close collaborations with EADS and Thales, within the Infom@gic project, and with INSERM and Mondaca within the ANR DAFOE project. The latter aims at developing better knowledge representations (ontologies) for satellite images and their applications to allow reasoning on these representations, using in particular the approaches described in Section [13.2.1]. The objective is to benefit from both the “bottom-up” learning approach and the “top-down” expert reasoning one, with applications to interactive satellite image annotation.

245
A new research axis concerns risk assessment and management, promoting the excellence of the research accomplished in the group by applying it to the management of disasters. These works concern the mapping of damages caused by disaster. The EXITER project was accomplished in collaboration with CNES as part of the international charter of risks. The EXITER project relies on the experience of the group in image analysis, knowledge extraction, classification and spatial reasoning.

13.2.6 Aerial and Satellite Imaging

**Faculty**  A. Almansa, J.-M. Nicolas, M. Roux, F. Tupin.

**Main events** Organization of the 2007 Urban Remote Sensing Joint Event (F. Tupin and M. Roux).


In aerial imaging, we extended our previous work to the processing of 3D point sets, acquired using laser techniques. Our contributions concern the soft non-parametric registration between such data and a numerical surface model obtained from photogrammetric images, in order to compensate for attitude movements of the sensor. Another contribution deals with change detection between two point sets or a point set and a polygonal model. These works are now integrated within the TerraNumerica project (CapDigital), which aims at modeling complete urban scenes via the fusion of aerial images and data acquired at ground level. Moreover, detection of objects in aerial and satellite images is addressed using learning methods based on Adaboost. Missing learning data were successfully compensated by generating examples through image synthesis. This work is carried out in collaboration with EADS.

Concerning Synthetic Aperture Radar imagery (SAR imagery), three main axes are developed. The first one is concerned with differential interferometry and ground movement monitoring, the second one deals with high resolution SAR imagery and optical and radar data fusion, and the third one with SAR data regularization.

In SAR differential interferometry, our works focus on two applicative and methodological fields. The first axis deals with subsidence study in Mexico in collaboration with the geology laboratory of ENS. The second axis is the glacier monitoring in the framework of MEGATOR project (ANR 2004-2007), which has lead to the development of a new SAR processor (SYTER) which is well adapted to high mountains. These two axes are now fused in a new project which started in 2008 for 4 years: EFIDIR (ANR MDCO). This project groups together 7 teams with methodologists and thematicians. All space agencies will provide SAR data in the framework of this project, specially of Argentiere glacier.

High resolution SAR imagery and the fusion of SAR and optical data is an important research axis, with increased interest due to the recent launch of metric SAR sensors in 2007 and 2008. A methodology of automatic registration has been developed, as well as a joint classification with SVM. In the frame of a CIFRE PhD with Thales, a processing chain for the detection of building and estimation of their height has been proposed. Interferometric aspects and 3D reconstruction have been studied in collaboration with ONERA and in a CNES project, and polarimetric aspects during the doctoral stay of Y. Wang (2008). The developments on SAR statistics and specially the Fisher distributions have been integrated in the active grid developed by Fresnel Institut. SAR urban areas have also been studied through a simulator of wave propagation.
The last axis deals with SAR data regularization. It is a recent research axis based on the development of two families of approaches: Markovian methods coupled with graph-cut optimization and non-local means. General contributions have been brought: first a fast graph-cut based algorithm for optimization of vectorial data have been developed; secondly, a probabilistic patch-based method has been proposed, which is able to deal with any kind of noise. These works have been applied to the regularization of amplitude data and interferometric data, specially in the frame of a CNES project and a collaboration of Naples University.

Other specific themes of SAR imagery have been developed. On change detection a collaboration has started with CEA in 2008. In the frame of a collaboration with Télécom Sud Paris in 2007, a classification coupling Fisher distributions and triplet Markov fields has been proposed. Improvements of previous works on road detection have been done in the frame of a collaboration with University of Pavie. A PhD on SAR data compression in relation with DGA has been led. Micro-Doppler have also been studied in a collaboration with ONERA.

In general, the team has developed an expertise on TerraSAR-X data through its participation to different projects, and specially on urban area processing. Moreover, its competence in coherent imagery (in particular on temporal approach) is used for sonar imagery (project with Telecom Bretagne) and in ultrasound imagery (PhD with SuperSonic Imagine).

### 13.3 References

#### 13.3.1 ACL: Articles in ISI-Indexed Journals


13.3. References


13. Image Processing and Interpretation (TII) 13.3. References


249
13.3.2 ACLN: Articles in Other Refereed Journals


13.3.3 INV: Invited Talks


13.3.4 ACTI: Articles in Proceedings of International Conferences


13. Image Processing and Interpretation (TI) 13.3. References


[2789] F. Cellier and E. Colin. Building height estimation using fine analysis of altimetric mixtures in layover areas on
13.3. References


[2791] F. Cellier, H. Oriot, and J. M. Nicolas. Hypothesis management for building reconstruction from high resolution InSAR imagery. In IGARSS'06, Denver, USA, August 2006.


[2820] M. Datcu, M. Soccorsi, D. Solimini, and F. Del Frate. Rate distortion theory method for eo data assessment and


[2874] V. Israel-Jost, E. Breton, E. Angelini, P. Choquet, I. Bloch, and A. Constantinescu. Vectorial Multi-Phase Mouse

[2873] C. Iorga and M. Datcu. Survey on methods for feature extraction of SAR data, ESA EUSC Conference on Information


13.3. References


13. Image Processing and Interpretation (TI)

13.3. References


13.3. References


13.3. References


13.3.5 ACTN: Articles in Proceedings of French Conferences


13.3.6 COM: Talks in Conferences Which Do Not Publish Proceedings


13.3.7 OS: Books and Book Chapters


261


### 13.3.8 AP-P: Patents


### 13.3.9 AP-R: Selected Technical Reports and Preprints


13.3. References