INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

# *Project-Team Magrit*

# *Visual Augmentation of Complex Environments*

*Nancy - Grand Est*

THEME COG

Activity Report

2008

# Table of contents

# 1. Team

**Research Scientist**
Marie-Odile Berger [ Research associate (CR) INRIA,Team Leader, HdR ]
Erwan Kerrien [ Research associate (CR) ]

**Faculty Member**
Gilles Simon [ Assistant professor, Université Henri Poincaré ]
Frédéric Sur [ Assistant professor, Institut National Polytechnique de Lorraine ]
Brigitte Wrobel-Dautcourt [ Assistant professor, Université Henri Poincaré ]

**External Collaborator**
René Anxionnat [ Medical Doctor, PhD, Professor CHRU Nancy ]

**Technical Staff**
Jonathan Ponroy [ Engineer, part-time work with the Parole team, until May 2008 ]
Frédéric Speisser [ Engineer, until August 2008 ]

**PhD Student**
Michael Aron [ INRIA, since October 2005 ]
Nicolas Noury [ INRIA since September 2006 ]
Nicolas Padoy [ ENS fellow, AMN, since October 2005 ]

**Post-Doctoral Fellow**
Evren Imre [ INRIA, since October 2007 ]

**Administrative Assistant**
Isabelle Herlich [ Research technician (TR) INRIA ]

# 2. Overall Objectives

## 2.1. Introduction

Augmented reality (AR) is a field of computer research which deals with the combination of real world and computer generated data in order to provide the user with a better understanding of his surrounding environment. Usually this refers to a system in which computer graphics are overlaid onto a live video picture or projected onto a transparent screen as in a head-up display.

Though there exist a few commercial examples demonstrating the effectiveness of the AR concept for certain applications, the state of the art in AR today is comparable to the early years of Virtual Reality. Many research ideas have been demonstrated but few have matured beyond lab-based prototypes.

Computer vision plays an important role in AR applications. Indeed, the seamless integration of computer generated objects at the right place according to the motion of the user needs automatic real-time detection and tracking. In addition, 3D reconstruction of the scene is needed to solve occlusions and light inter-reflexion between objects and to make easier the interactions of the user with the augmented scene. Since fifteen years, much work has been successfully devoted to the problem of structure and motion, but these works are often formulated as off-line algorithms and require batch processing of several images acquired in a sequence. The challenge is now to design robust solutions to these problems with the aim to let the user free of his motion during AR applications and to widen the range of AR application to large and/or unstructured environments. More specifically, the Magrit team aims at addressing the following problems:

- On-line pose computation for structured and non structured environments: this problem is the cornerstone of AR systems and must be achieved in real time with a good accuracy.
- Long term management of AR applications: a key problem of numerous algorithms is the gradual drifting of the localization over time. One of our aims is to develop methods that improve the accuracy and the repeatability of the pose during arbitrarily long periods of motion.

- 3D modeling for AR applications: this problem is fundamental to manage light interactions between real and virtual objects, to solve occlusions and to obtain realistic fused images.

## 2.2. Highlights

Sébastien Gorges received the Second Prize of the Région Lorraine in February 2008 for his thesis about the use of augmented reality for interventional neuronavigation. This work has been achieved in close collaboration with the CHU Nancy Hospital and GE Healthcare.

# 3. Scientific Foundations

## 3.1. Scientific Foundations

The aim of the Magrit project is to develop vision based methods which allow significant progress of AR technologies in terms of ease of implementation, usability, reliability and robustness in order to widen the current application field of AR and to improve the freedom of the user during applications. Our main research directions concern two crucial issues, camera tracking and scene modeling. Methods are developed with a view to meet the expected robustness and to provide the user with a good perception of the augmented scene.

### *3.1.1. Camera calibration and registration*

**Keywords:** *Registration*, *augmented reality*, *tracking*, *viewpoint computation*.

One of the most basic problems currently limiting Augmented Reality applications is the registration problem. The objects in the real and virtual worlds must be properly aligned with respect to each other, or the illusion that the two worlds coexist will be compromised.

As a large number of potential AR applications are interactive, real time pose computation is required. Although the registration problem has received a lot of attention in the computer vision community, the problem of real-time registration is still far from being a solved problem, especially for unstructured environments. Ideally, an AR system should work in all environments, without the need to prepare the scene ahead of time, and the user should walk anywhere he pleases.

For several years, the Magrit project has aimed at developing on-line and markerless methods for camera pose computation. Within the European Project ARIS, we have proposed a real-time system for camera tracking designed for indoor scenes. The main difficulty with online tracking is to ensure robustness of the process. Indeed, for off-line processes, robustness is achieved by using spatial and temporal coherence of the considered sequence through move-matching techniques. To get robustness for open-loop systems, we have developed a method which combines the advantage of move-matching methods and model-based methods [6] by using a piecewise-planar model of the environment. This methodology can then be used in a wide variety of environments: indoor scenes, urban scenes ... We are also concerned with the development of methods for camera stabilization. Indeed, statistical fluctuations in the viewpoint computations lead to unpleasant jittering or sliding effects, especially when the camera motion is small. We have proved that the use of model selection allows us to noticeably improve the visual impression and to reduce drift over time.

An important way to improve the reliability and the robustness of pose algorithms is to combine the camera with another form of sensor in order to compensate for the shortcomings of each technology. Indeed, each technology approach has limitations: on the one hand, rapid head motions cause image features to undergo large motion between frame that can cause visual tracking to fail. On the other hand, inertial sensors response is largely independent from the user's motion but their accuracy is bad and their response is sensible to metallic objects in the scene. We have proposed a system that makes an inertial sensor (MT9- Xsens) cooperate with the camera based system in order to improve the robustness of the AR system to abrupt motions of the users, especially head motion. This work contributes to reduce the constraints on the users and the need to carefully control the environment during an AR application [1]. This research area is continued within the ASPI project in order to build a dynamic articulatory model from various image modalities and sensor data.

It must be noted that the registration problem must be addressed from the rather specific point of view of augmented reality: the success and the acceptation of an AR application does not only depend on the accuracy of the pose computation but also on the visual impression of the augmented scene. The search for the best compromise between accuracy and perception is therefore an important issue in this project. This research topic is currently addressed in our project both in classical AR [7] and in medical imaging in order to choose the camera model, including intrinsic parameters, which describes at best the considered camera.

Finally, camera tracking largely depends on the quality of the matching stage which allows to detect and to match features over the sequence. Ongoing research are conducted on the problem of establishing robust correspondences of features over time. The use of *a contrario* decision is currently under study to achieve this aim [5].

### 3.1.2. Scene modeling

**Keywords:** *Fusion*, *medical imaging*, *reconstruction*, *volumetric segmentation*.

Modeling the scene is a fundamental issue in AR for many reasons. First, pose computation algorithms often use a model of the scene or at least some 3D knowledge on the scene. Second, effective AR systems require a model of the scene to support occlusion and to compute light reflexions between the real and the virtual objects. Unlike pose computation which has to be computed in a sequential way, scene modeling can be considered as an off-line or an on-line problem according to the application.

In the Magrit team, scene modeling was mainly addressed as an off-line and possibly interactive process, especially to build models for medical imaging from several images modalities. Since two years, one of our research directions is about online scene reconstruction, with the aim to be able to handle AR applications in vast environments without the need to instrument the scene.

Interactive scene modeling from various image modalities is mainly considered in our medical activities. For the last 15 years, we have been working in close collaboration with the neuroradiology laboratory (CHU-University Hospital of Nancy) and GE Healthcare. As several imaging modalities are now available in a per-operative context (2D and 3D angiography, MRI, ...), our aim is to develop a multi-modality framework to help therapeutic decision and treatment.

In previous works [4], we proposed an efficient solution to the registration of 2D/3D angiographic images and 3DXA/MRI images. Since then, we have mainly been interested in the effective use of a multimodality framework in the treatment of arteriovenous malformations (AVM). The treatment of AVM is classically a two-stage process: embolization or endovascular treatment is first performed. This step is then followed by a stereotactic irradiation of the remnant. Hence an accurate definition of the target is a parameter of great importance for the treatment. Our short term aim is to perform an accurate detection of the AVM shape within a multimodality framework. Our long term aim is to develop multimodality and augmented reality tools which make cooperate various image modalities (2D and 3D angiography, fluoroscopic images, MRI, ...) in order to help physicians in clinical routine. One of the success of this collaboration is the implementation of the concept of *augmented fluoroscopy* [3], which helps the surgeon to guide endoscopic tools towards the malformation.

Multi-modality techniques for reconstruction are also considered within the european ASPI project, the aim of which is to build a dynamic model of the vocal tract from various images modalities (MRI, ultrasound, video) and magnetic sensors.

Besides interactive modeling, research on on-line reconstruction are conducted in our team. Sequential reconstruction of the scene structure needed by pose or occlusion algorithms is highly desirable for numerous AR applications for which instrumentation is not conceivable. Hence, structure and pose must be sequentially estimated over time. We are currently studying this problem both for multi-planar scenes and for general scenes.

# 4. Application Domains

## 4.1. Augmented reality

**Keywords:** *Multimedia.*

We have a significant experience in the AR field especially through the European project ARIS (2001–2004) which aimed at developing effective and realistic AR systems for e-commerce and especially for interior design. Beyond this restrictive application field, this project allowed us to develop nearly real time camera tracking methods for multi-planar environments. Since then, we have amplified our research on multi-planar environments in order to obtain effective and robust AR systems in such environments. We especially aim at automatically building a model of the scene during the applications in order to be able to consider large and unknown environments.

## 4.2. Medical imaging

**Keywords:** *Health.*

For 15 years, we have been working in close collaboration with the University hospital of Nancy and GE Healthcare in interventional neuroradiology with the aim to develop tools allowing the physicians to take advantage of the various existing imaging modalities on the brain in their clinical practice. As several imaging modalities that bring complementary information on the various brain pathologies are now available in a pre-operative context (subtracted angiography 2D and 3D, fluoroscopy, MRI,...) our aim is to develop a multi-modality framework to help therapeutic decisions. Recently, we have investigated the use of AR tools for neuronavigation. The aim of the PhD thesis of Sebastien Gorges which ended in May 2007 was to design tools for neuronavigation that take advantage of a real-time imagery (fluoroscopy) and a pre-operative 3D imagery (3D angiography). We are currently involved in the ARC SIMPLE with the ALCOVE team and the University hospital of Nancy. Our aim is to develop simulation tools of the interventional act adapted to the patient's anatomy and physiology, in order to help the surgeon with planning the coil placement, rehearsing the therapeutic gesture, and to provide new tools to improve the medical training to the technique.

## 4.3. Talking Head

**Keywords:** *Multimedia.*

We are involved in the FET-STREP european ASPI project which started on November 2005. There is a strong evidence that visual information of the speaker, especially jaws and lips, noticeably improves the speech intelligibility. Hence, having a realistic talking head could help language learning technology in giving the student a feedback on how to change articulation in order to achieve a correct pronunciation. This task is complex and necessitates a multidisciplinary effort involving speech production modeling and image analysis. The long term aim of the ASPI project is the design of a 3D +t articulatory model to be used for the realistic animation of a talking head. Within this project, we especially work on the tracking of the visible articulators using stereo-vision techniques and we intend to supplement the model with internal articulators (tongue, larynx) obtained from medical imaging (ultrasound images for tongue tracking and MRI for global model).

# 5. Software

## 5.1. RAlib

Our software efforts are integrated in a library called RAlib which contains our research development on image processing, registration (2D and 3D) and visualization. This library is licensed by the APP (French agency for software protection). The visualization module is called QGLSG: it enables the visualization of images, 2D and 3D objects under a consistent perspective projection. It is based on Qt (http://www.trolltech.com) and OpenScenegraph (http://www.openscenegraph.org/projects/osg) libraries. The latter was integrated in the project by Frédéric Speisser, who joined the Magrit project-team in September 2006 as an INRIA assistant engineer. Before Frédéric Speisser left us at the end of July 2008, he consolidated the integration of OpenScenegraph features (picking, shadows, texture mapping) and finalized the doxygen documentation. New features were also developed grounding on OpenScenegraph shader functionalities: online image distortion correction is now available for video camera flows, invisible objects can also be incorporated in a scene so that virtual objects can cast shadows on real objects, and occlusion between virtual and real objects are easier to handle. The library was also ported to Mac OS and Windows. A great effort was also made this year to disseminate the library within the team, and many full-size applications were written to sustain all our research projects.

## 5.2. Patents

Our collaboration with GE Healthcare has given rise to several patent disclosures on specific calibration process, registration and visualization [20].

# 6. New Results

## 6.1. New Results

### *6.1.1. Scene and camera reconstruction*

**Participants:** Marie-Odile Berger, Nicolas Noury, Gilles Simon, Frédéric Sur, Evren Imre.

On the theme of scene and camera reconstruction, we followed two main directions of research, one in which the objects of the scenes are planes, the other in which the environment is unspecified. In this latter case, probabilistic frameworks were considered in order to meet the required robustness.

#### *6.1.1.1. Structure from motion via a contrario models*

Structure from motion problems call for probabilistic frameworks to meet robustness requirements. Features (e.g. points) are extracted from images, then they are matched under projective constraints. This determines both the structure of the scene and the position of the camera. The problem is difficult since the position of the features may not be accurately known, and matching may introduce false correspondences which can endanger the reconstruction process. We aim at developing new probabilistic methods to tackle these problems. Publication [16] focuses on fundamental matrix estimation, which is a keystep in many structure from motion problems. Fundamental matrix estimation is actually based on correspondences that are spoilt by noise and outliers. Outliers must be thrown out via robust statistics, and noise gives uncertainty. In [16] we provide a closed-form formula for the uncertainty of the so-called 8 point algorithm, which is a basic tool for fundamental matrix estimation via robust methods. As an application, we modify a well established *a contrario* robust algorithm accordingly, leading to a new criterion to recover point correspondences under epipolar constraint, balanced by the uncertainty of the estimation.

#### *6.1.1.2. Online reconstruction for AR tasks*

While integrating user cues is very common in off-line computer vision, user survey has almost never been used in a simultaneous localization and mapping (SLAM) context. One reason is that designating features is much more complicated in a flow of images acquired at runtime than in a static image or a pre-acquired sequence. Moreover, the automatic frameworks used to simultaneously build a map and localize the camera generally preclude any knowledge insertion during the full-parameter computation stage.

To encompass these problem and benefit from user expertise, we decompose the global estimation problem into several simpler ones, allowing safer resolution of each separate task as well as automatic or user-guided assessments of the intermediate results [18]. As an example, intersection lines between the world planes are detected using a time-consistent filtering technique. These lines, which are intermediate results toward the Euclidean reconstruction of the scene, can be easily assessed by the user by simple image comparison.

This year, we improved the robustness of the system by adding a recovery procedure based on keyframes and SIFT features [15]. We are also working on handling more complex primitives in the scene such as cubes and other standard geometric objects. The goal is to build what we could call an "immersive modeling tool" optimized for augmented reality applications.

#### *6.1.1.3. Improved inverse-depth parameterization for SLAM*

Inverse-depth parameterization [21] is an approach that successfully deals with the feature initialization problem in monocular simultaneous localization and mapping applications. However, it is computationally costly, and when a set of landmarks is initialized from the same image, it fails to enforce the common origin constraint. We thus propose two improvements of the classical inverse-depth parameterization. The key-idea is to factor out the common pose parameters when several landmarks are initialized from the same image. In the first extension (IDP1), only the pose is factored out whereas in the second one (IDP2), both the pose and the orientation are factored out. Experiments proved that IDP2 is superior to the classical IDP both in computational cost and in performance, whereas IDP1 delivers a similar performance at a much lower computational cost.

### *6.1.2. Medical imaging*

**Participants:** René Anxionnat, Marie-Odile Berger, Erwan Kerrien, Nicolas Padoy, Cédric Laurent.

#### *6.1.2.1. Simulation for planning the embolization of intracranial aneurisms*

The endovascular treatment for an intracranial aneurism consists in filling the aneurismal cavity by placing coils. These are sorts of long platinum springs that, once deployed, wind into a compact ball. Considering the location of the lesion, close to the brain, and its small size, a few millimeters, the interventional gesture requires a good planning and can only be performed by a very experienced surgeon. A simulation tool of the interventional act, available in the operating room, reliable, adapted to the patient's anatomy and physiology, would help to plan the coil placement, rehearse the procedure, and improve the medical training to the technique. The SIMPLE project, an INRIA cooperative research action (ARC), started in 2007 and ended this year. It aimed at developing methods to simulate coil deployment in an intracranial aneurism, running in real time and adaptable to any patient data. We coordinated this project which ran in collaboration with the Alcove project-team at INRIA Lille-Nord Europe and the Department of Interventional Neuroradiology at University Hospital of Nancy. Our task consists in providing a precise arterial geometric model as well as information on blood flow specific to the patient. The first step was to extract the arterial wall. Despite the very high quality of the available 3D images (3D rotational angiography), tomographic reconstruction artefacts perturb the isosurface that should correspond to the arterial wall. Taking this isosurface as an initialization, we proposed to improve it within an active surface framework where the arterial wall is deformed until its X-ray projection fits a set of registered 2D angiographic images taken on the patient. Our algorithm was finalized this year and published in [13]. The second step addressed the blood flow: what has to be modeled and how to model it? Our contribution consisted in devising image acquisition protocols and image processing algorithms to measure blood flow characteristics in-vivo. Cédric Laurent joined the team for a 3-month internship as an engineer in computational fluid dynamics (CFD). He set up a suite of free CFD software called Gerris Flow

Solver (Gfs available at http://gfs.sf.net) and found a 4% error on average on simulated data for the algorithm he wrote to measure the blood flow velocity. No ground truth could be established for in-vivo measurements using the algorithm. However, they were compatible with US doppler measurements. Besides, the Alcove project-team was in charge of physically modeling the coil. We participated in the model validation in a constraint-free environment. This joint work was published in [11]. The joint effort was pursued, with the result that coil simulation in real patient data can now be performed. Validation data were acquired both on a silicon phantom and 2 patients at Nancy University Hospital. We processed the data and validation experiments are currently being conducted.

*6.1.2.2. Surgical workflow analysis*

Surgery rooms are complex environments where many interactions take place between staff members and the electronic and mechanical systems. In spite of their inherent complexity, surgeries of the same kind bear numerous similarities and are usually performed with similar workflows. This gives the possibility to design support systems in the Operating Room (OR), whose applicability range from easy tasks such as the activation of OR lights and calling the next patient, to more complex ones such as context-sensitive user interfaces or automatic reporting. An essential feature when designing such systems, is the ability for on-line recognition of what is happening inside the OR, based on recorded signals. In [14], we presented an approach using signals from the OR and Hidden Markov Models to recognize on-line the surgical steps performed by the surgeon during a laparoscopic surgery. We also explained how the system can be deployed in the OR. Experiments were presented using 11 real surgeries performed by different surgeons in several ORs, recorded at our partner hospital. We believe that similar systems will quickly develop in the near future in order to efficiently support surgeons, trainees and the medical staff in general, as well as to improve administrative tasks like scheduling within hospitals.

### 6.1.3. Modeling face dynamics

**Participants:** Michael Aron, Marie-Odile Berger, Erwan Kerrien, Jonathan Ponroy, Brigitte Wrobel-Dautcourt.

Being able to produce realistic facial animation is crucial for many speech applications in language learning technologies. In order to reach realism, it is necessary to acquire 3D models of the face and of the internal articulators (tongue, palate,...) from various image modalities.

*6.1.3.1. Modeling the vocal tract*

Our long term objective is to provide intuitive and near-automatic tools for building a dynamic 3D model of the vocal tract from various image and sensor modalities (MRI, ultrasound (US), video, magnetic sensors ...). Previous works have proven that the ultrasound modality was an efficient way to acquire dynamic tongue information. Unfortunately, the tip of the tongue is not visible in most US images because air stops the propagation of the ultrasounds. In this work, we use magnetic sensors that are glued on the tongue in order to complete the tongue contour obtained from ultrasound images near the apex.

Combining several modalities requires that all geometrical and temporal data be consistent together. It also requires to define appropriate image processing techniques to extract the articulators (tongue, palate, lips...) from the data. During the previous year, a fast, low cost and easily reproducible acquisition system has been designed in order to temporally align the data. During this year, we especially focussed on the processing of the US articulatory data which bring important information on the tongue dynamics.

Tracking the tongue contour in US sequences is difficult for many reasons: (i) the high speed of the tongue for some sounds produces blurred contours which are difficult to track, (ii) the tongue undergoes highly elastic deformations (iii) the apex is often not visible as the air of the sub-lingual cavity prevents sound propagation. The tracker we have designed [10] operates in a loop following two major stages: a prediction of the contour is computed under the hypothesis of a parametric model. A refinement stage based on active contour models is then performed to compute the actual position. Adaptation of this general scheme to cope with the specificities of the tongue and to improve the robustness of the tracking includes the following: (i) a specific anisotropic filtering has been designed to enhance the tongue contour (ii) appropriate boundary conditions for the active models are used (iii) robustness of the tracking is enforced by introducing priors on the tongue shape provided by the EM sensors glued on the tongue.

This method was successfully tested on speech sequences even when fast motions occurred. Residual failures were due to difficulties in imaging the back of the tongue. An interface incorporating this tracking tool was designed to recover from possible failures. This allowed us to extract easily the tongue shape in the vast amount of US sequences recorded in our corpus.

Finally, we addressed the problem of fusing image modalities. As 3D measures of the face can be extracted from both the MRI and the stereocopic images, MRI and videos were registered thanks to the use of an iterative closest point algorithm. All the modalities were then registered using some EM sensors glued on the US probe and under the ears of the speaker. As a result, dynamic articulatory including points on the lips, the tongue and the palate are now available [9]. These data have been very recently used to perform articulatory inversion. To the best of our knowledge, this is the first work that demonstrates the potential of static and dynamic data fusion in the construction of articulatory databases.

*6.1.3.2. Realistic face animation*

Modeling facial dynamics is often achieved using stereovision techniques. Unfortunately, reconstruction artifacts are common and spatial and temporal regularization constraints are often imposed to cope with this problem. As for most regularization methods, the parameters of the regularization functions need to be carefully tuned to obtain satisfying results in order to avoid reconstruction artifacts or excessive smoothing.

In order to improve the robustness of the process, we proposed an approach that borrowed concepts from model-based approach and from vision based tracking of markers on the face. Our input data is a reduced set of 3D dense maps of the talker obtained for some sustained sounds as well as a corpus of stereo sequences of the talker with markers painted on his/her face that allowed the face kinematics to be computed. In this study, the 3D dense maps were acquired with a scanner for a sustained vowel but other acquisition technologies can be used. Our main idea was to transfer kinematics learned on the sparse mesh onto the 3D dense mesh in order to generate realistic dense animations of the face. During this year we especially focused on the recovering of the dense modes of the talker, thus allowing an easy animation of the head. These dense modes were computed from the sparse ones, assuming that the coefficients of the dense modes are identical to the coefficient of the sparse modes. Preliminary experiments proved the effectiveness of the method.

# 7. Contracts and Grants with Industry

## 7.1. Partnerships

### *7.1.1. GE Healthcare*

The partnership with GE Healthcare (formerly GE Medical Systems) started in 1993. In the past few years, it bore on the supervision of CIFRE PhD fellows on the topic of using a multi-modal framework in interventional neuroradiology. The concept of *Augmented Fluoroscopy* was one of the main results of this PhD. A prototype that implements these results has been developed by GE Healthcare and has been available at the University Hospital of Nancy for clinical evaluation since July 2006.

# 8. Other Grants and Activities

## 8.1. National Initiatives

### *8.1.1. ARC SIMPLE*

**Participants:** René Anxionnat, Marie-Odile Berger, Erwan Kerrien, Cédric Laurent.

The SIMPLE project, an INRIA cooperative research action (ARC), started in 2007 and ended this year. It aimed at developing methods to simulate coil deployment in an intracranial aneurism, running in real time and adaptable to any patient data. We coordinated this project which ran in collaboration with the Alcove project-team at INRIA Lille-Nord Europe and the Department of Interventional Neuroradiology at University Hospital of Nancy. Last year seminal work was consolidated this year, leading to 2 publications. The first [11] at Medical Image Computing and Computer Assisted Intervention conference (MICCAI: the main conference in medical imaging) reported on the coil model and its validation in a constraint-free environment. The second [13] at Augmented Medical Environment workshop (AMI-ARCS: one of MICCAI satellite workshops), described our surface refinement algorithm to retrieve a very precise and smooth surface for the blood vessels. Our work on modeling the vasculature was pursued by the development of image processing algorithms to measure the blood flow velocity in-vivo. Besides, data were acquired and processed both on a silicon phantom and 2 patients as a basis for the currently on-going validation of the coil deployment simulation. Back on vessel geometric modeling, we are currently investigating the use of implicit surface models within our segmentation framework, since such models are more adapted to the subsequent simulation.

## 8.2. European initiatives

### *8.2.1. ASPI–IST FET STREP*

**Participants:** Michael Aron, Marie-Odile Berger, Nicolas Ferveur, Erwan Kerrien, Brigitte Wrobel-Dautcourt.

ASPI is about Audiovisual-to-articulatory inversion. Participants in this project are INRIA Nancy Grand Est, ENST (Paris), KTH (Stokholm), the University Research Institute of National Technical University of Athens and the University of Bruxelles. Audiovisual-to-articulatory inversion consists in recovering the vocal tract shape dynamics (from vocal folds to lips) from the acoustical speech signal, supplemented by image analysis of the speaker's face. Being able to recover this information automatically would be a major break-through in speech research and technology, as a vocal tract representation of a speech signal would be both beneficial from a theoretical point of view and practically useful in many speech processing applications (language learning, automatic speech processing, speech coding, speech therapy, film industry...). The Magrit team is involved in the development of articulatory models from various image modalities (ultrasound, video, MRI) and electromagnetic sensors.

This year, the first main achievement concerned the design of automatic processes for extracting the tongue contour of the US images. The second contribution is the fusion of features extracted -lips, tongue, palate - from the various image modalities. The reliability of these dynamic articulatory data has been accessed using acoustic-to-articulatory inversion techniques [12].

## 8.3. International initiatives

### *8.3.1. Conferences, meetings and tutorial organization*

- M.-O. Berger was a member of the program committee of the conferences ISMAR 08, ICPR 08 and of the workshops AMI-ARCS and NORDIA (Non-Rigid Shape Analysis and Deformable Image Alignment).
- G. Simon was a member of the program committee of ISMAR 08.
- E. Kerrien was a member of the program committee of MICCAI 08.

# 9. Dissemination

## 9.1. Teaching

- Several members of the group, in particular assistant professors and Ph.D. students, actively teach at Henri Poincaré Nancy 1, Nancy 2 universities and INPL.
- Other members of the group also teach in the computer science Master of Nancy and in the “Master en sciences de la vie et de la santé” (SVS).

## 9.2. Participation to conferences and workshops

Members of the group participated in the following events:

International Conference on Medical Image Computing and Computer Assisted Intervention (New York), British Machine Vision Conference (Leeds), International Conference on Pattern Recognition (Tampa, Florida), International Seminar on Speech Production (Strasbourg), European Signal Processing Conference (Lausanne), Conference on Innovative Applications of Artificial Intelligence (Chicago), Workshop AMI-ARCS (New York), Reconnaissance des Formes et Intelligence artificielle (Amiens).

## 9.3. Participation to PhD thesis committees

The members of the team took part to the following thesis committees:

- M.O. Berger: Geraldo Silveira (Ecole des Mines), Mathias Fontmarty (LAAS), Michael Baumann (TIMC Grenoble), Christophe Marecaux (Grenoble), Jonathan Boisvert (Nice/Montreal).

# 10. Bibliography

## Major publications by the team in recent years

[1] M. Aron, G. Simon, M.-O. Berger. *Use of Inertial Sensors to Support Video Tracking*, in "Computer Animation and Virtual Worlds", vol. 18, 2007, p. 57-68, http://hal.inria.fr/inria-00110628/en/.

[2] M.-O. Berger, R. Anxionnat, E. Kerrien, L. Picard, M. Soderman. *A methodology for validating a 3D imaging modality for brain AVM delineation: Application to 3DRA.*, in "Computerized Medical Imaging and Graphics", vol. 32, 2008, p. 544-553, http://hal.inria.fr/inria-00321688/en/.

[3] S. Gorges, E. Kerrien, M.-O. Berger, Y. Trousset, J. Pescatore, R. Anxionnat, L. Picard, S. Bracard. *3D Augmented Fluoroscopy in Interventional Neuroradiology: Precision Assessment and First Evaluation on Clinical Cases*, in "Workshop on Augmented environments for Medical Imaging and Computer-aided Surgery - AMI-ARCS 2006 (held in conjunction with MICCAI'06), Copenhagen Danemark", Wolfgang Birkfellner, Nassir Navab and Stephane Nicolau, 11 2006, http://hal.inria.fr/inria-00110850/en/.

[4] E. Kerrien, M.-O. Berger, E. Maurincomme, L. Launay, R. Vaillant, L. Picard. *Fully automatic 3D/2D subtracted angiography registration*, in "Proceedings of Medical Image Computing and Computer-Assisted Intervention - MICCAI'99, Cambridge, England", C. Taylor, A. Colchester (editors), Lecture Notes in Computer Science, vol. 1679, Springer, 1999, p. 664-671.

[5] P. Musé, F. Sur, F. Cao, Y. Gousseau, J.-M. Morel. *An a contrario decision method for shape element recognition*, in "International Journal of Computer Vision", vol. 69, n$^{o}$ 3, 2006, p. 295-315, http://hal.inria.fr/inria-00104260/en/.

[6] G. Simon, M.-O. Berger. *Pose estimation for planar structure*, in "IEEE Computer Graphics and Applications", vol. 22, n$^{o}$ 6, 2002, p. 46-53.

[7] J.-F. VIGUERAS, G. SIMON, M.-O. BERGER. *Calibration Errors in Augmented Reality: A Practical Study*, in "International Symposium on Mixed and Augmented Reality, Vienna Autriche", 10 2005, p. 154-163, http://hal.inria.fr/inria-00000383/en/.

## Year Publications

### Articles in International Peer-Reviewed Journal

[8] M.-O. BERGER, R. ANXIONNAT, E. KERRIEN, L. PICARD, M. SÖDERMAN. *A methodology for validating a 3D imaging modality for brain AVM delineation: Application to 3DRA.*, in "Computerized Medical Imaging and Graphics", vol. 32, 2008, p. 544-553, http://hal.inria.fr/inria-00321688/en/.

### International Peer-Reviewed Conference/Proceedings

[9] M. ARON, M.-O. BERGER, E. KERRIEN. *Multimodal fusion of electromagnetic, ultrasound and MRI data for building an articulatory model*, in "International Seminar On Speech Production, France Strasbourg", 2008, http://hal.inria.fr/inria-00326290/en/.

[10] M. ARON, A. ROUSSOS, M.-O. BERGER, E. KERRIEN, P. MARAGOS. *Multimodality Acquisition of Articulatory Data and Processing*, in "EUSIPCO, Suisse Lausanne", EPFL, 2008, http://hal.inria.fr/inria-00316041/en/.

[11] J. DEQUIDT, M. MARCHAL, C. DURIEZ, E. KERRIEN, S. COTIN. *Interactive Simulation of Embolization Coils: Modeling and Experimental Validation*, in "Medical Imaging Computing and Computer Assisted Intervention, MICCAI Lecture Notes in Computer Science, États-Unis d'Amérique New York City", vol. 5241, 2008, p. 695-702, http://hal.inria.fr/inria-00336907/en/.

[12] A. KATSAMANIS, A. ROUSSOS, P. MARAGOS, M. ARON, M.-O. BERGER. *Inversion from Audiovisual Speech to Articulatory Information by Exploiting Multimodal Data*, in "International Seminar On Speech Production, France Strasbourg", 2008, http://hal.inria.fr/inria-00327031/en/.

[13] E. KERRIEN, M.-O. BERGER, J. DEQUIDT. *Refining the 3D surface of blood vessels from a reduced set of 2D DSA images*, in "AMI-ARCS 2008, États-Unis d'Amérique New York, NY", PJ "Eddie" Edwards and Kensaku Mori and Tobias Sielhorst, 2008, p. 61-69, http://hal.inria.fr/inria-00321550/en/.

[14] N. PADOY, B. TOBIAS, H. FEUSSNER, M.-O. BERGER, N. NAVAB. *On-line Recognition of Surgical Activity for Monitoring in the Operating Room*, in "Proceedings of the 20th Conference on Innovative Applications of Artificial Intelligence,, États-Unis d'Amérique Chicago", 2008, p. 1718-1724, http://hal.inria.fr/inria-00331390/en/.

[15] G. SIMON, M.-O. BERGER. *Detection of the Intersection Lines in Multiplanar Environments: Application to Real-Time Estimation of the Camera-Scene Geometry*, in "International Conference on Pattern Recognition (ICPR 2008), États-Unis d'Amérique Tampa", 2008, http://hal.inria.fr/inria-00322127/en/.

[16] F. SUR, N. NOURY, M.-O. BERGER. *Computing the uncertainty of the 8 point algorithm for fundamental matrix estimation*, in "British Machine Vision Conference, Royaume-Uni Leeds", 2008, http://hal.inria.fr/inria-00319704/en/.

[17] M. TODA, S. MAEDA, M. ARON, M.-O. BERGER. *Modeling Subject-Specific Formant Transition Patterns in /aSa/ Sequences*, in "International Seminar On Speech Production, France Strasbourg", 2008, http://hal.inria.fr/inria-00326295/en/.

### National Peer-Reviewed Conference/Proceedings

[18] G. SIMON, M.-O. BERGER. *Reconstruction et augmentation simultanées de scènes planes par morceaux*, in "16e congrès francophone AFRIF-AFIA Reconnaissance des Formes et Intelligence Artificielle - RFIA 2008, France Amiens", 2008, http://hal.inria.fr/inria-00186303/en/.

### Scientific Books (or Scientific Book chapters)

[19] F. CAO, J. L. LISANI, J.-M. MOREL, P. MUSÉ, F. SUR. , J.-M. MOREL, F. TAKENS, B. TEISSIER (editors)*A Theory of Shape Identification*, Springer, 2008, http://hal.inria.fr/inria-00319719/en/.

### Other Publications

[20] S. GORGES, Y. TROUSSET, J. PESCATORE, V. BISMUTH, M.-O. BERGER, E. KERRIEN. *System and method to improve visibility of an object in an imaged subject*, Extension to Europe of US patent US20070740734 published on 2007-04-26 G06T17/40; G06T17/40, 10 2008, http://hal.inria.fr/inria-00336786/en/, patent no. EP1986159.

## References in notes

[21] J. MONTIEL, J. CIVERA, A. DAVISON. *Unified Inverse Depth Parametrization for Monocular SLAM*, in "Robotics: Science and Systems", July 2006, http://pubs.doc.ic.ac.uk/inverse-depth-slam/.