INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team magrit

Visual Augmentation of Complex Environments

Nancy - Grand Est

Theme : Vision, Perception and Multimedia Understanding

Activity Report

2010

Table of contents

1. Team

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

Faculty Members
Gilles Simon [Assistant professor, Université Henri Poincaré]
Frédéric Sur [Assistant professor, Institut National Polytechnique de Lorraine]
Pierre-Frédéric Villard [Assistant professor, Université Henri Poincaré, since September 2009]
Brigitte Wrobel-Dautcourt [Assistant professor, Université Henri Poincaré]

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

Technical Staff
Christel Leonet [Engineer, from October 2010]

PhD Students
Michael Aron [INRIA, Until February 2010]
Srikrishna Bhat [INRIA, since December 2008]
Nicolas Noury [INRIA since September 2006]
Nicolas Padoy [Until April 2010]
Abdulkadir Eryildirim [CNRS, since January 2010]
Ahmed Yureidini [INRIA, since January 2010]

Administrative Assistant
Isabelle Herlich [INRIA]

Other
Pierre Escamilla [Student intern]

2. Overall Objectives

2.1. Overall Objectives

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.

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. Scientific Foundations

3.1. Camera calibration and registration

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 been aiming at developing on-line and markerless methods for camera pose computation. We have especially proposed a real-time system for camera tracking designed for indoor scenes [1]. The main difficulty with online tracking is to ensure robustness of the process. 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 by using a piecewise-planar model of the environment. This methodology can 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. Each technology approach has limitations: on the one hand, rapid head motions cause image features to undergo large motion between frames 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 sensitive to metallic objects in the scene. We have proposed a system that makes an inertial sensor cooperate with the camera-based system in order to improve the robustness of the AR system to abrupt motions of the users, especially head motions. 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 has been continued within the ASPI project in order to build a dynamic articulatory model from various image modalities and sensor data.

Obtaining a model of the scene where the AR applications is to take place is often required by pose algorithms. However, obtaining a model either by automatic or interactive means is a tedious task, especially for large environments. In addition, models may be described in terms of 3D features which cannot be identified in the images. Pose by recognition is thus an appealing approach which allows to link photometric knowledge learned on the scene to the camera pose. We are currently considering learning-based techniques, the aim of which is to allow pose computation from video sequences previously acquired on the site where the application is to be used.

Finally, it must be noted that the registration problem must be addressed from the specific point of view of augmented reality: the success and the acceptance 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 has been addressed in our project both in classical AR and in medical imaging in order to choose the camera model, including intrinsic parameters, which describes at best the considered camera.

3.2. Scene modeling

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 our past activities, 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 three 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.

We have mainly been interested in the effective use of a multimodality framework in the treatment of arteriovenous malformations (AVM) and aneurysms in the context of interventional neuroradiology. The goal of interventional gestures is to guide endoscopic tools towards the pathology with the aim to perform embolization of the AVM or to fill the aneurysmal cavity by placing coils. An accurate definition of the target is a parameter of great importance for the success of the treatment. We have proposed and developed 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*, which helps the surgeon to guide endoscopic tools towards the pathology. Lately, in cooperation with the Shaman EPI, we have proposed new methods for implicit modeling of the aneurysms with the aim to obtain near real time simulation of the coil deployment in the aneurysm [3]. 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.

On-line 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. This process largely depends on the quality of the matching stage which allows to detect and to match features over the sequence. Ongoing research are thus conducted on the use of probabilistic methods to establish robust correspondences of features over time. The use of *a contrario* decision is especially under study to achieve this aim [4].

Most automatic techniques aim at reconstructing a sparse and thus unstructured set of points of the scene. Such models are obviously not appropriate to perform interaction with the scene. In addition, they are incomplete in the sense that they may omit features which are important for the accuracy of the pose recovered from 2D/3D correspondences. We have thus investigated interactive techniques with the aim to obtain reliable and structured models of the scene. The goal of our approach is to develop immersive and intuitive interaction techniques which allow scene modeling during the application [15].

4. Application Domains

4.1. Augmented reality

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 currently investigate both automatic and interactive techniques for scene reconstruction/structure from motion methods in order to be able to consider large and unknown environments.

4.2. Medical Imaging

For 15 years, we have been working in close collaboration with University Hospital of Nancy and GE Healthcare in interventional neuroradiology. Our common aim is to develop a multimodality framework to help therapeutic decisions and interventional gestures. In particular, we aim at developing tools allowing the physicians to take advantage of the various existing imaging modalities on the brain in their clinical practice: 2D subtracted angiography (2DSA), 3D rotational angiography (3DRA), fluoroscopy, MRI,...Recent works concern the use of AR tools for neuronavigation and the development of simulation tools of the interventional act for training or planning. This last project is developed in collaboration with the EPI Shaman.

4.3. Augmented head

Visual information on a speaker, especially jaws and lips but also tongue position, noticeably improves speech intelligibility. Hence, having a realistic augmented head displaying both external and internal articulators could help language learning technology progress in giving the student a feedback on how to change articulation in order to achieve a correct pronunciation. The long term aim of the project is the acquisition of articulatory data and the design of a 3D+t articulatory model from various image modalities: external articulators are extracted from stereovision data, the tongue shape is acquired through ultrasound imaging, 3D images of all articulators can be obtained with MRI for sustained sounds, magnetic sensors are used to recover the tip of the tongue.

5. Software

5.1. Software

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[1] and OpenScenegraph[2] libraries. The QGLSG library integrates innovative features such as online camera distortion correction, and invisible objects that can 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 and a full doxygen documentation was written.

[1] http://www.trolltech.com
[2] http://www.openscenegraph.org/projects/osg

6. New Results

6.1. Scene and camera reconstruction

Participants: Marie-Odile Berger, Srikrishna Bhat, Nicolas Noury, Gilles Simon, Frédéric Sur.

On the theme of scene and camera reconstruction, we investigate both fully automatic methods and learning-based techniques for pose and structure recovery. Interactive techniques are also considered in order to obtain well-structured description of the scene and to meet the required robustness with the help of the user.

6.1.1. A statistical framework for robust image point matching

Matching interest points is the cornerstone of many computer vision applications. Our contributions to this field are twofold: (i) developing robust matching algorithms which take account of the uncertainty of point location and (ii) building up matching algorithms which are robust to wide viewpoint changes and to repeated patterns. These two main lines are based on statistical models (namely *a contrario* models).

This year, we have investigated [16], [23] a general *a contrario* algorithm to match features by taking account of point uncertainty, both for 2D (SIFT keypoints) or 3D (data from depth cameras). Applications of this framework to SIFT matching and 3D data fusion have been developped and proved the effectiveness of the method.

Making matching algorithms robust to repeated patterns is of primary importance in man-made environments. We have proposed a new *a contrario* matching algorithm incorporating both descriptor similarity and geometric constraints [19], [22] in a new one-stage approach.

This method is robust to repeated patterns and improves the number of retrieved correspondences. However, it still suffers from a lack of robustness with respect to strong viewpoint changes. Viewpoint simulation is a recent and powerful technique to achieve viewpoint invariance (see e.g. ASIFT [26]). The idea behind ASIFT is to simulate image views obtainable by varying the camera axis orientation. We recently improved ASIFT by incorporating the aforementioned *a contrario* robust matching [12]. This algorithm, called I-ASIFT (Improved ASIFT), is robust to strong viewpoint change and to repeated patterns. From this point of view, it performs better than any existing point matching algorithm.

6.1.2. Learning-based techniques for pose computation

Learning based approaches attempt to quantize the descriptors extracted from a set of images of the environment into clusters, called visual words. Representing an image by a vector of word frequencies has proven to be an efficient way of matching and is used for relocalization of lost cameras in AR systems. We investigate the use of learning-based techniques for pose computation in S. Bhat's PhD thesis. Our aim is to build visual words dedicated to pose computation. We have designed words for pose computation with the idea to only retain clusters which are likely to represent one physical 3D point. The rationale behind this was to reduce the potential set of 2D/3D matching hypothesis especially when large environments are considered, making RANSAC-based pose algorithms more tractable. Our visual words were constructed using transitive closure techniques from a sequence taken continuously with a moving camera [10]. The interest of these words is that no prior constraints were imposed on the shape of the words, contrary to the classical K-means approach. Thereby, common problematic situations can largely be prevented, such as where words may contain features related to different physical 3D objects, or where features corresponding to one physical 3D point are split across several words.

6.1.3. Online reconstruction for AR tasks

Our motivation is to enable fast capture of textured surface models of environments suited to AR. Methods like SFM, SLAM as well as laser scans generate 3D point clouds that have to be manually post-treated to get the desired description. By contrast, in-situ modeling allows a user to directly build a surface model of his/her surrounding environment and verify the geometry against the physical world in real-time. The recovered 3D model can be integrated on-the-fly in an AR loop or stored for future use. This year, we have proposed a novel method for in-situ 3D sketching of polyhedral scenes. The underlying idea is to transpose single view reconstruction techniques to an immersive context [15]. Interactions were inspired by those described in [27] and used in Google SketchUp™. The main difference resided in that we considered dynamic video images instead of static ones. For instance, moving a mouse cursor over a fixed image was replaced by moving an image (rotating the camera) behind a fixed cursor (e.g. at center of the view). As a global result, we now have an easy-to-use and intuitive modeling method particularly adapted to mouse-less mobile devices such as tablet PC, PDA and mobile phones.

6.2. Medical imaging

Participants: René Anxionnat, Marie-Odile Berger, Pierre Escamilla, Erwan Kerrien, Nicolas Padoy, Pierre-Frédéric Villard, Ahmed Yureidini.

6.2.1. Realistic patient-based simulators for interventional gestures

6.2.1.1. Interventional Neuroradiology

Our research activity is led in collaboration with Shaman project-team from INRIA Lille-Nord Europe and the Department of Interventional Neuroradiology from Nancy University Hospital. It was pursued this year in the context of the SOFA-InterMedS INRIA Large-Scale Initiative. A. Yureidini started his PhD thesis under the shared direction of Magrit and Shaman INRIA project-teams in January 2010, on that topic, following P. Glanc's engineering internship.

Our objective is the modeling of blood vessels from 3DRA data, with the aim to use these models for real time simulation of interventional procedures. Implicit models were favored for their adaptation to simulation. In particular blobby models [28] were under investigation this year. The following issues were addressed:

- robust vessel tracking to both initialize the location of seminal blobs for the model and extract points on the vessel surface from 3DRA data. An original algorithm was designed, relying on successive robust, RANSAC-based, fitting of cylinders along the vessels.
- methods to interpolate unstructured points using an implicit model, by minimizing an energy. We observed that replacing the classical algebraic distance by an approximation of the geometric distance [25] in the data attachment term improved the behavior of the dynamic system under optimization. In particular, a new criterion to detect worst approximating blobs was derived, leading to an efficient blob splitting mechanism to further refine the model.

The resulting blobby model was implemented in Sofa platform and a first live demonstration of the simulation was presented during Sofa-InterMeds evaluation seminar in last October, showing an impressive realism during tool navigation.

Implicit models are not naturally able to handle specific topological situations, and tend to blend in two close vessels that should remain separate in the simulation. A new model consisting of a tree of local implicit models was developed, were each cylinder detected in the tracking phase is replaced by a local implicit submodel. Methods to integrate such models in the simulation were described, leading to a specific surface extraction procedure.

The full model is currently under evaluation and in the process of being published.

6.2.1.2. Abdominal procedures

The goal here is to help interventionnal radiology trainees in learning the crucial tasks of the abdominal procedures with a computer-based simulator. One of the advantages of such medical simulators include being able to practice on different anatomies with different pathologies. However one of the key challenges is to correctly reproduce the physiological behavior such as breathing.

In collaboration with Imperial College of London, University of California, Bangor University and INSA of Lyon with have proposed a medical simulator for training [29]. It computes directly on the GPU the X-ray attenuation from polygon meshes dynamically modified depending on the respiration cycle of the virtual patient. The simulations were however restricted to monochromatic X-ray beams and finite point sources.

We have extended this year in [21] our previous simulation pipeline to take into account focal spots that cause geometric unsharpness and polychromatic X-rays. Dynamic polygon meshes can still be used as input data of the respiration model.

6.2.2. Image-guided simulation based on a biomedical model of the respiratory motion

This work has been done in collaboration with GE Healthcare. The problematic is to augment abdominal dynamic fluoroscopy with 3D motion information due to the respiration. The goal is to determine the organ positions and then decrease the contrast agent dose used by the physician to locate his/her surgical tools.

Our strategy is to 1) extract the patient's anatomy from an initial CT scan; 2) elaborate a respiration model such as the one studied in collaboration with the University Claude Bernard Lyon 1 [14] and 3) monitor the model with the fluoroscopy.

We focused this year on the rib kinematics during a M.Sc. internship. It is based on rotation centers computed for each ribs from the initial CT scan and rotation angles computed with fluoroscopy contour analysis. We have also used SOFA (Simulation Open-Framework Architecture) to test our initial fully-thoracic respiration model and the resulting simulation achieved real-time.

6.3. Modeling face and vocal tract dynamics

Participants: Michael Aron, Marie-Odile Berger, Abdulkadir Eryildirim, Erwan Kerrien, 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.

6.3.1. shape-based variational framework for curve segmentation

Segmenting the vocal tract in MRI is difficult especially because the tongue may move near other edges in the oral cavity, such as the palate or the teeth, which may disturb the segmentation process. The idea explored in our work is to guide the segmentation with shape priors learnt on a reference speaker within a a shape-based variational framework. Following the work in [24], we have proposed a total energy including both global and local image statistics. The global term extracts roughly the object in the whole image domain while the local term improves precision inside a small neighborhood around the contour. Shape priors are incorporated into segmentation via a PCA model with a relatively large number of components to enable the adaptation of the model to strong morphological differences. Results are really promising [13]. We are currently investigating how to build models respecting physical correspondences from this purely segmentation step. Correlation-based techniques have been recently developped to automatically extract physically-corresponding start and end points of the tongue. These constraints have been incorporated within the variational framework thus allowing to recover corresponding tongue contours.

6.3.2. Acquisition of articulatory data

Our aim is to acquire data on a speaker's vocal tract during locution and to build models from these data. The acquisition system was progressively built up during the three ellapsed years. Compared to previously existing systems, where many assumptions are made concerning sensor accuracy and/or spatial registration and/or temporal synchronization, our system produces well founded data. During this year, we especially focus on the evaluation of the uncertainty attached to the data provided by the system. The accuracy of each system unit, let it be either a modality, or a registration or synchronization procedure, was experimentally measured. Finally, Monte-Carlo sinmulations were carried out to infer the statistical uncertainty on the whole system as well as every sub-system [20]. As a result, we are the first to propose a certified system to acquire data on the vocal tract. These studies also gave preliminary insights on how to further improve the quality of the system. In particular, we designed a phantom dedicated to the calibration of the US/EM registration.

6.4. National Initiatives

- SOFA-InterMedS (2009–) Participants: R. Anxionnat, M.O. Berger, E. Kerrien, P. Glanc, A. Yureidini.
  The SOFA-InterMedS large-scale INRIA initiative[3] is a research-oriented collaboration across several INRIA project-teams, international research groups and clinical partners. Its main objective is to leverage specific competences available in each team to further develop the multidisciplinary field of Medical Simulation research. Our action within the initiative takes place in close collaboration with both Shaman INRIA project-team in Lille and the Department of diagnostic and therapeutic interventional neuroradiology of Nancy University Hospital. We aim at providing in-vivo models of the patient's organs, and in particular a precise geometric model of the arterial wall. Such a model is used by Shaman team to simulate the coil deployment within an intracranial aneurysm. The associated medical team in Nancy, and in particular our external collaborator René Anxionnat, is in charge of validating our results.
- ANR ARTIS (2009-2012)
  Participants: M.O. Berger, A. Eryildirim, E. Kerrien.
  The main objective of this fundamental research project is to develop inversion tools and to design and implement methods that allow for the production of augmented speech from the speech sound signal alone or with video images of the speaker's face. The Magrit team is especially concerned with the development of procedures allowing the automatic construction of a speaker's model from various imaging modalities.
- ANR Visac (2009-2012)
  Participants: M.O. Berger, B. Wrobel-Dautcourt.
  The ANR Visac is about acoustic-visual speech synthesis by bimodal concatenation. The major challenge of this project is to perform speech synthesis with its acoustic and visible components simultaneously. Within this project, the role of the Magrit team is twofold. One of them is to build a stereovision system able to record synchronized audio-visual sequences at a high frame rate. Second, a highly realistic dense animation of the head must be produced. During this year, a setup for Acoustic-Visual Speech Synthesis by Concatenating Bimodal Units has been designed [17].

7. Dissemination

7.1. Animation of the scientific community

- The members of the team frequently review articles and papers for IEEE TBME (Trans. Biomed. Eng.), IJCARS (International Journal of Computer Assisted Radiology and Surgery), TVCG (IEEE Transactions on Visualization and Computer Graphics), IEEE Transactions on Multimedia, CAVW (Computer Animation and Virtual Worlds), EURASIP Journal on Image and Video Processing, IEEE Transactions on Image Processing, Signal Image and Video Processing, SIAM Journal on Imaging Sciences, EAAI (Engineering Applications of Artificial Intelligence), Progress in Biophysics and Molecular Biology, AIMS (Inverse Problems and Imaging).
- M.O. Berger was a member of the program committee of the following conferences: International Symposium on Mixed and Augmented Reality (ISMAR 2010); International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2010); International Conference on Pattern Recognition (ICPR 2010); the IEEE International Symposium on Biomedical Imaging (ISBI) 2011. She was also a member of the program committee of the French conference on pattern recognition (RFIA 2010).

[3] http://en.inria.fr/research/research-fields/large-scale-initiatives

- E. Kerrien was a member of the program committee of: MICCAI 2010; and the African Conference on Research in Computer Science and Applied Mathematics (CARI) 2010.
- G. Simon was a member of the program committee of: ISMAR 2010, Virtual Reality International Conference(VRIC 2010).
- F. Sur is a member of the editorial board of the journal Image Processing On Line. He was a member of the program committee of CARI 2010.
- P.F. Villard was a member of the program committee of: the International Symposium on Biomedical Simulation (ISBMS) 2010: the IEEE International Symposium on Biomedical Imaging (ISBI) 2010, ICPR 2010, MICCAI 2010.

7.2. 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 ”DIU Chirurgie Robotique ”.

7.3. Participation to conferences and workshops

Members of the group participated in the following events: International Conference on Acoustics, Speech, and Signal Processing (Dallas), International Conference on Pattern Recognition (Istanbul), International Symposium on Visual Computing (Las Vegas), Eurographics 2010 (Norrköping, Sweden), ULTRAFEST conference 2010 (New Haven), RFIA 2010 (Caen).

8. 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", 2007, vol. 18, 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", 2008, vol. 32, p. 544-553, http://hal.inria.fr/inria-00321688/en/.

[3] 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", USA, 2008, vol. 5241, p. 695-702, http://hal.inria.fr/inria-00336907/en/.

[4] 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", 2006, vol. 69, n^{o} 3, p. 295-315 [*DOI :* 10.1007/S11263-006-7546-0], http://hal.inria.fr/inria-00104260/en/.

[5] N. NOURY, F. SUR, M.-O. BERGER. *How to Overcome Perceptual Aliasing in ASIFT?*, in "6th International Symposium on Visual Computing - ISVC 2010", Etats-Unis Las Vegas, Lecture Notes in Computer Science, Springer, Sep 2010, http://hal.inria.fr/inria-00515375/en.

[6] N. Padoy, T. Blum, A. Ahmadi, H. Feussner, M.-O. Berger, N. Navab. *Statistical Modeling and Recognition of Surgical Workflow*, in "Medical Image Analysis", 2010, To be published.

[7] G. Simon. *In-Situ 3D Sketching Using a Video Camera as an Interaction and Tracking Device*, in "31st Annual Conference of the European Association for Computer Graphics - Eurographics 2010", SuÃ¨de Norrkoping, May 2010, http://hal.inria.fr/inria-00474324/en.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[8] N. Padoy. *Modélisation des Activités Chirurgicales et de leur Déroulement pour la Reconnaissance des Etapes Opératoires*, Université Henri Poincaré - Nancy I, Apr 2010, http://hal.inria.fr/tel-00487069/en.

Articles in International Peer-Reviewed Journal

[9] N. Padoy, T. Blum, A. Ahmadi, H. Feussner, M.-O. Berger, N. Navab. *Statistical Modeling and Recognition of Surgical Workflow*, in "Medical Image Analysis", 2010, To be published.

International Peer-Reviewed Conference/Proceedings

[10] S. Bhat, M.-O. Berger, G. Simon, F. Sur. *Transitive Closure based visual words for point matching in video sequence*, in "20th International Conference on Pattern Recognition - ICPR 2010", Turquie Istanbul, 2010, http://hal.inria.fr/inria-00486749/en.

[11] U. Musti, A. Toutios, S. Ouni, V. Colotte, B. Wrobel-Dautcourt, M.-O. Berger. *HMM-based Automatic Visual Speech Segmentation Using Facial Data*, in "Interspeech 2010", Japon Makuhari, Chiba, ISCA, Sep 2010, p. 1401-1404, http://hal.inria.fr/inria-00526776/en.

[12] N. Noury, F. Sur, M.-O. Berger. *How to Overcome Perceptual Aliasing in ASIFT?*, in "6th International Symposium on Visual Computing - ISVC 2010", Etats-Unis Las Vegas, Lecture Notes in Computer Science, Springer, Sep 2010, http://hal.inria.fr/inria-00515375/en.

[13] T. Peng, E. Kerrien, M.-O. Berger. *A shape base framework to segmentation of tongue contours from MRI data*, in "35th IEEE International Conference on Acoustics, Speech, and Signal Processing - ICASSP 2010", États-Unis Dallas, IEEE, Mar 2010, p. 662 - 665 [*DOI :* 10.1109/ICASSP.2010.5495123], http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5495123, http://hal.inria.fr/inria-00442138/en.

[14] J. Saadé, A.-L. Didier, P.-F. Villard, R. Buttin, J.-M. Moreau, M. Beuve, B. Shariat. *A Preliminary Study For A Biomechanical Model Of The Respiratory System*, in "Engineering and Computational Sciences for Medical Imaging in Oncology - ECSMIO 2010", France Angers, May 2010, 7, Engineering and Computational Sciences for Medical Imaging in Oncology - ECSMIO is the special session 1 of International Conference on Computer Vision Theory and Applications - VISAPP 2010, http://hal.inria.fr/hal-00509817/en.

[15] G. Simon. *In-Situ 3D Sketching Using a Video Camera as an Interaction and Tracking Device*, in "31st Annual Conference of the European Association for Computer Graphics - Eurographics 2010", Suède Norrköping, May 2010, http://hal.inria.fr/inria-00474324/en.

[16] F. Sur. *Robust matching in an uncertain world*, in "20th International Conference on Pattern Recognition - ICPR 2010", Turquie Istanbul, IEEE, Aug 2010, p. 2350 - 2353, ISSN: 1051-4651
 Print

ISBN: 978-1-4244-7542-1 [*DOI* : 10.1109/ICPR.2010.575], http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5597003, http://hal.inria.fr/inria-00474083/en.

[17] A. TOUTIOS, U. MUSTI, S. OUNI, V. COLOTTE, B. WROBEL-DAUTCOURT, M.-O. BERGER. *Setup for Acoustic-Visual Speech Synthesis by Concatenating Bimodal Units*, in "Interspeech 2010", Japon Makuhari, Chiba, ISCA, Sep 2010, p. 486-489, http://hal.inria.fr/inria-00526766/en.

[18] A. TOUTIOS, U. MUSTI, S. OUNI, V. COLOTTE, B. WROBEL-DAUTCOURT, M.-O. BERGER. *Towards a True Acoustic-Visual Speech Synthesis*, in "9th International Conference on Auditory-Visual Speech Processing - AVSP2010", Japon Hakone, Kanagawa, Sep 2010, p. POS1-8, http://hal.inria.fr/inria-00526782/en.

National Peer-Reviewed Conference/Proceedings

[19] N. NOURY, F. SUR, M.-O. BERGER. *Modèle a contrario pour la mise en correspondance robuste sous contraintes épipolaires et photométriques*, in "17ième congrès francophone AFRIF-AFIA, Reconnaissance des Formes et Intelligence Artificielle - RFIA 2010", France Caen, Université de Caen Basse-Normandie and laboratoire GREYC, Jan 2010, http://hal.inria.fr/inria-00432992/en.

Workshops without Proceedings

[20] M. ARON, M.-O. BERGER, E. KERRIEN. *Evaluation of the uncertainty of multimodal articulatory data*, in "Ultrafest V", États-Unis New Heaven, Doug Whalen, Khalil Iskarous, Aude Noiray, Mar 2010, http://hal.inria.fr/inria-00429330/en.

[21] F. VIDAL, P.-F. VILLARD, M. GARNIER, N. FREUD, J. LÉTANG, N. JOHN, F. BELLO. *Joint Simulation of Transmission X-ray Imaging on GPU and Patient's Respiration on CPU*, in "American Association of Physicists in Medicine Annual Meeting - AAPM 2010", États-Unis Philadelphia, Jul 2010, http://hal.inria.fr/hal-00485720/en.

Research Reports

[22] N. NOURY, F. SUR, M.-O. BERGER. *Determining point correspondences between two views under geometric constraint and photometric consistency*, INRIA, Apr 2010, n^{o} RR-7246, http://hal.inria.fr/inria-00471874/en.

[23] F. SUR. *Robust matching in an uncertain world*, INRIA, Sep 2010, n^{o} RR-7374, http://hal.inria.fr/inria-00515974/en.

References in notes

[24] T. F. CHAN, L. A. VESE. *Active contours without edges*, in "IEEE Trans. on Image Processing", 2001, vol. 10, n^{o} 2.

[25] GABRIEL TAUBIN. *Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation*, in "IEEE Trans. PAMI", November 1991, vol. 13, n^{o} 11, p. 1115-1138.

[26] J.-M. MOREL, G. YU. *A new framework for fully affine invariant image comparison*, in "SIAM Journal on Imaging Sciences", 2009, vol. 2, n^{o} 2, p. 438–469.

[27] J. OH, W. STUERZLINGER, J. DANAHY. *Comparing SESAME and sketching on paper for conceptual 3D design*, in "EUROGRAPHICS Workshop 2005", 2005.

[28] SHIGERU MURAKI. *Volumetric shape description of range data using "blobby model"*, in "Computer Graphics", July 1991, vol. 25, n^{o} 4, p. 227-235.

[29] P.-F. VILLARD, F. VIDAL, C. HUNT, F. BELLO, W. J. NIGEL, S. JOHNSON, D. GOULD. *A prototype percutaneous transhepatic cholangiography training simulator with real-time breathing motion*, in "International Journal of Computer Assisted Radiology and Surgery", 11 2009, vol. 4, n^{o} 6, p. 571-578.