INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team LAGADIC

Visual servoing in robotics, computer vision, and augmented reality

Rennes

THEME COG

Table of contents

1. Team

Head of project-team
François Chaumette [DR Inria]

Administrative assistant
Stéphanie Lemaile [TR Inria, shared with Aspi, Imadoc, and Visages teams]

Inria research scientists
Christophe Collewet [Cemagref research scientist, on Inria secondment (détachement) from 01/10/2005]
Alexandre Krupa [CR]
Éric Marchand [CR]

Inria technical staff
Fabien Spindler [IR]

Teaching assistant
Anthony Remazeilles [Insa, University of Rennes 1]

Inria Post-doctoral fellows
Sinisa Segvic [Mobivip project, from 01/07/2005]

Inria project technical staff
Cédric Riou [Sora project, until 31/03/2005]
Fabien Dionnet [ESA/Trasys project, from 15/11/2005]

Inria junior technical staff
Anne-Sophie Tranchant [shared with Visages team, until 30/04/2005]

Ph-D students
Odile Bourquardez [DGA/CNRS grant]
Andrew Comport [Inria grant until 31/03/2005]
Claire Dune [Inria-Brittany council grant from 01/11/2005]
Nicolas Mansard [ENS Cachan-Bretagne member until 31/08/2005, Inria grant from 01/09/2005]
Muriel Pressigout [Research Ministry grant]
Romeo Tatsambon [Research Ministry grant from 01/11/2005]

2. Overall Objectives

2.1. Overall Objectives

Keywords: *active vision*, *augmented reality*, *computer vision*, *image sequence*, *robot vision*, *tracking*, *visual servoing*.

Research activities of the Lagadic team are concerned with visual servoing and active vision. Visual servoing consists in using the information provided by a vision sensor to control the movements of a dynamic system, this system which can be real within the framework of robotics, or virtual within the framework of computer animation or augmented reality. This research topic is at the intersection of the fields of robotics, automatic control, and computer vision. These fields are the subject of profitable research since many years and are particularly interesting by their very broad scientific and application spectrum. Within this spectrum, we focus ourselves on the interaction between visual perception and action. This topic is significant because it provides an alternative to the traditional Perception-Decision-Action cycle. It is indeed possible to link more closely the perception and action aspects, by directly integrating the measurements provided by a vision sensor in closed loop control laws.

This set of themes of visual servoing is the central scientific topic of the Lagadic group. More generally, our objective is to design strategies of coupling perception and action from images for applications in robotics, computer vision, virtual reality and augmented reality.

This objective is significant, first of all because of the variety and the great number of the potential applications to which can lead our work. It is also significant to be able to raise the scientific aspects associated with these problems, namely modeling of visual features representing in an optimal way the interaction between action and perception, taking into account of complex environments and the specification of high level tasks. We also work to treat new problems provided by imagery systems such those resulting from an omnidirectional vision sensor or echographic probes. We are finally interested in revisiting traditional problems in computer vision (3D localization, structure and motion) through the visual servoing approach.

3. Scientific Foundations

3.1. Visual servoing

Basically, visual servoing techniques consist in using the data provided by one or several cameras in order to control the motions of a dynamic system [37] [5]. Such systems are usually robot arms, or mobile robots, but can also be virtual robots, or even a virtual camera. A large variety of positioning tasks, or mobile target tracking, can be implemented by controlling from one to all the degrees of freedom of the system. Whatever the sensor configuration, which can vary from one on-board camera on the robot end-effector to several free-standing cameras, a set of visual features has to be selected at best from the image measurements available, allowing to control the degrees of freedom desired. A control law has also to be designed so that these visual features $\mathbf{s}(t)$ reach a desired value $\mathbf{s}^*$, defining a correct realization of the task. A desired trajectory $\mathbf{s}^*(t)$ can also be tracked. The control principle is thus to regulate to zero the error vector $\mathbf{s}(t) - \mathbf{s}^*(t)$. With a vision sensor providing 2D measurements, potential visual features are numerous, since as well 2D data (coordinates of feature points in the image, moments, ...) as 3D data provided by a localization algorithm exploiting the extracted 2D features can be considered. It is also possible to combine 2D and 3D visual features to take the advantages of each approach while avoiding their respective drawbacks [6].

More precisely, a set $\mathbf{s}$ of k visual features can be taken into account in a visual servoing scheme if it can be written:

$$\mathbf{s} = \mathbf{s}(\mathbf{x}(\mathbf{p}(t)), \mathbf{a}) \tag{1}$$

where $\mathbf{p}(t)$ describes the pose at the instant t between the camera frame and the target frame, $\mathbf{x}$ the image measurements, and $\mathbf{a}$ a set of parameters encoding a potential additional knowledge, if available (such as for instance a coarse approximation of the camera calibration parameters, or the 3D model of the target in some cases).

The time variation of $\mathbf{s}$ can be linked to the relative instantaneous velocity $\mathbf{v}$ between the camera and the scene:

$$\dot{\mathbf{s}} = \frac{\partial \mathbf{s}}{\partial \mathbf{p}}\, \dot{\mathbf{p}} = \mathbf{L_s}\, \mathbf{v} \tag{2}$$

where $\mathbf{L_s}$ is the interaction matrix related to $\mathbf{s}$. This interaction matrix plays an essential role. Indeed, if we consider for instance an eye-in-hand system and the camera velocity as input of the robot controller, we obtain when the control law is designed to try to obtain an exponential decoupled decrease of the error:

$$\mathbf{v}_c = -\lambda \hat{\mathbf{L_s}}^+ (\mathbf{s} - \mathbf{s}^*) - \hat{\mathbf{L_s}}^+ \frac{\hat{\partial \mathbf{s}}}{\partial t}$$

where λ is a proportional gain that has to be tuned to minimize the time-to-convergence, $\hat{\mathbf{L_s}}^+$ is the pseudo-inverse of a model or an approximation of the interaction matrix, and $\frac{\hat{\partial \mathbf{s}}}{\partial t}$ an estimation of the target velocity.

From the selected visual features and the corresponding interaction matrix, the behavior of the system will have particular properties as for stability, robustness with respect to noise or to calibration errors, robot 3D trajectory, etc. Usually, the interaction matrix is composed of highly non linear terms and does not present any decoupling properties. This is generally the case when $\mathbf{s}$ is directly chosen as $\mathbf{x}$. In some cases, it may lead to inadequate robot trajectories or even motions impossible to realize, local minimum, tasks singularities, etc. [2]. It is thus extremely important to "cook" adequate visual features for each robot task or application, the ideal case (very difficult to obtain) being when the corresponding interaction matrix is constant, leading to a simple linear control system. To conclude in few words, **visual servoing is basically a non linear control problem. Our Graal quest is to transform it as a linear control problem.**

Furthermore, embedding visual servoing in the task function approach [39] allows to solve efficiently the redundancy problems that appear when the visual task does not constrain all the degrees of freedom of the system. It is then possible to realize simultaneously the visual task and secondary tasks such as visual inspection, or joint limits or singularities avoidance. This formalism can also be used for tasks sequencing purposes.

4. Application Domains

4.1. Application Domains

Keywords: *augmented reality*, *image-guided neuro-surgery*, *robotics*, *vehicle navigation*.

The natural applications of our research are obviously in robotics. In the past, we mainly worked in the field of the grasping and of the manipulation of tools, in the field of underwater robotics for the stabilization of images, and the positioning of uninstrumented robot arms, in the field of agro-industry for the positioning of a vision sensor in order to ensure an improvement of the quality controls of agro-alimentary products, as well as in the field of the video surveillance (control of the movements of a pan-tilt camera to track mobile natural objects). More recently, we addressed the field of mobile robotics via the activities undertaken around the Cycab vehicle: detection and tracking of mobile objects (pedestrians, other vehicles), control by visual servoing of the movements of the vehicle.

In fact, researches which we undertake in the Lagadic group can apply to all the fields of robotics implying a vision sensor. They are indeed conceived to be independent of the robot system considered (the robot and the vision sensor can even be virtual as we will see it in the continuation).

Currently, we are interested in using visual servoing for the control of robot arms in space, and underactuated flying robots, such as miniature helicopters and aircrafts.

In collaboration with the Visages team, we also address the field of medical robotics. The applications under consideration for the moment turn around new functionalities of assistance to the clinician during a medical examination: visual servoing on echographic images, active perception for the optimal generation of 3D echographic images, coupling between an off-set vision of the field of examination and a force sensor for tele-operated examinations, etc.

Robotics is not the only possible application field to our researches. In the past, we were interested in collaboration with the Siames project to apply the techniques of visual servoing in the field of computer animation. It can be a question either of controlling the movement of virtual humanoids according to their pseudo-perception, or to control the point of view of visual restitution of an animation. In both cases, potential applications are in the field of virtual reality, for example for the realization of video games, or virtual cinematography.

Applications also exist in computer vision and augmented reality. It is then a question of carrying out a virtual visual servoing for the 3D localization of a tool with respect to the vision sensor, or for the estimation of its 3D motion. This field of application is very promising, because in full rise for the realization of special effects in the multi-media field or for the design and the inspection of objects manufactured in the industrial world.

Lastly, our work in visual servoing and active perception can be related with those carried out in cogni-science, in particular in the field of psychovision (for example on the study of eye motion in the animal and human visual system, or on the study of the representation of perception, or on the study of the links between action and perception).

5. Software

5.1. ViSP2 : a visual servoing platform

Participants: Éric Marchand, Fabien Spindler, François Chaumette.

Visual servoing is a very active research area in vision-based robotics. A software environment that allows fast prototyping of visual servoing tasks is then of prime interest. The main reason is certainly that it usually requires specific hardware (the robot and, most of the time, dedicated image framegrabbers). The consequence is that the resulting applications are often not portable and cannot be easily adapted to other environments. Today's software design allows one to propose elementary components that can be combined to build portable high-level applications. Furthermore, the increasing speed of micro-processors allows the development of real-time image processing algorithms on a usual workstation. We have developed a library of canonical vision-based tasks for eye-in-hand and eye-to-hand visual servoing that contains the most classical linkages that are used in practice. The ViSP software environment features all the following capabilities : independence with respect to the hardware, simplicity, extendibility, portability. Moreover, ViSP involves a large library of elementary positioning tasks w.r.t. various basic visual features (points, lines, circles, spheres, cylinders,...) that can be combined together, and an image processing library that allows the tracking of visual cues (dot, segment, ellipse, spline,...). A simulation software has also been developed.

This year, this modular platform has been fully rewritten (leading to ViSP2) and has been released as a free software under the QPL licence. ViSP has been primarily developed in C++ on Unix workstations and is now available on Linux and Mac OSX. The new version of ViSP and its full functionalities are described in [18]. It can be downloaded from `http://www.irisa.fr/lagadic/visp`.

5.2. Marker software : Marker-based augmented reality kernel

Participant: Éric Marchand.

The Marker software implements an algorithm supplying the computation of camera pose and camera calibration using fiducial markers. The parameters estimation is handled using virtual visual servoing. The principle consists in considering the pose and the calibration as a dual problem of visual servoing. This method presents many advantages: similar accuracy as for the usual non-linear minimization methods, simplicity, effectiveness. A licence of this software was yielded to the Total Immersion company.

5.3. MarkerLess : MarkerLess-based augmented reality kernel

Participants: Cédric Riou, Andrew Comport, Éric Marchand.

Markerless is an upgrade of the Marker software with additional features developed within the SORA Riam Project. It allows the computation of camera pose with no fiducial marker.

A real-time, robust and efficient 3D model-based tracking algorithm for a monocular vision system has been developed. Tracking objects in the scene amounts to computing the pose between the camera and the objects. Non-linear pose computation is formulated by means of a virtual visual servoing approach. In this context, the derivation of point-to-curves interaction matrices are given for different features including lines, circles, cylinders and spheres. A local moving-edge tracker is used in order to provide a real-time estimation of the displacements normal to the object contours. A method is proposed for combining local position uncertainty and global pose uncertainty in an efficient and accurate way by propagating uncertainty. Robustness is obtained by integrating an M-estimator into the visual control law via an iteratively re-weighted least squares implementation. More recently, we also considered the case of non-rigid objects. The proposed method has been validated on several complex image sequences including outdoor environments. Application for this tracker is robotics and visual servoing as well as augmented reality applications.

5.4. Development work: Robot vision platforms

Participant: Fabien Spindler.

We exploit several experimental platforms to validate our research work in visual servoing and in active vision. More precisely, we have two robotic systems built by Afma Robots. The first one is a cartesian robot with six degrees of freedom, the other one is a cylindrical robot with four degrees of freedom. Each robot is equipped with a CCD camera mounted on the effector and an Imaging Technology framegrabber. This year, we have installed two additional cameras on the cartesian robot: a Marlin firewire color camera and an other CCD camera equipped with a wireless video transmission system. A PC on Linux communicates with the robot using a SBS Technologies bus adapter. These equipments require specific hardware, but also software maintenance actions and new developments in order to make them evolve. Training and assistance of the users, presentation of demonstrations also form part of the daily activities. This year, all the drivers (cameras, framegrabbers, robots) and demonstrations were adapted to use the new version of our software visual servoing platform ViSP2 (see Section 5.1). Moreover, a six minutes video was achieved to explain our activities in robotics [36].

5.5. Development work : Medical robot

Participants: Anne-Sophie Tranchant, Fabien Spindler, Alexandre Krupa.

Fall 2004, our new robotic system equipped with an ultrasound probe and dedicated to medical applications was declared opened to service by the Sinters company. The beginning of this year was devoted to additional tests of the system. Based on new developments made in order to improve the robustness of the robot, an eye to hand visual servoing application was first designed to validate the real-time control loop of the robot at 50 fps. Secondly, an automatic probe calibration was built using adaptive ultrasound image-based visual servoing (see Section 6.2.1 and [21]). The calibration parameters are than exploited to reconstruct 3D echographies.

5.6. Development work : Cycab

Participants: Fabien Spindler, Nicolas Mansard.

The Cycab is a small autonomous electric car dedicated to vision-based mobile robotics applications. This year, many hardware and software improvements were brought to simplify the developments. Considering the material, the reducers were replaced to have more torque to transport two passengers on a sloping road. In order to suppress all the image processing difficulties due to reflections on the windshield, a new pan-tilt firewire camera (Biclops PTM) was installed on the front avoid-shock. Considering the software, the network configuration was modified to authorize the control of the Cycab from a laptop. Moreover resulting from a collaboration with Inria Grenoble, a simulator and a 3D visualizer engine were installed. By the end of the year, new developments should make it possible to control the vehicle directly through the simulator, the transformation of a simulated application to a real application becoming thus immediate as for the control part.

6. New Results

6.1. Visual servoing

6.1.1. Geometric visual features modeling

Participants: Romeo Tatsambon, François Chaumette.

This study is directly related to the search of optimal visual features, as described in Section 3.1. Following our recent works on selecting adequate combination of image moments as visual features for visual servoing [1] [33], we have considered vanishing points, that can be computed in the image from the perspective projection of parallel 3D straight lines. Vanishing points are well known in the computer vision community to be invariant with respect to any 3D translational motion. This property is particularly interesting in visual servoing to design visual features able to control the rotational motions. Different parameterizations have been studied, mainly cartesian coordinates, and cylindrical coordinates. The experimental results obtained on a dedicated target (a rectangle) have validated the nice properties of the control scheme built from such features (large convergence domain, decoupling properties, adequate robot trajectories), as it was expected from the theoretical analysis. Finally, fall 2005, we have started to consider affine invariants as potential visual features.

6.1.2. Photometric features modeling

Participants: Christophe Collewet, François Chaumette.

One of the main problem in visual servoing is to extract and track robustly the image measurements that are used to build the visual features involved in the control scheme. This may lead to complex and time consuming image processing, and may increase the effect of image noise. To cope with this problem, we propose to use directly photometric features, as the luminance for example, as input of the control scheme. As described in Section 3.1, the main challenge is to obtain the interaction matrix related to such features. Theoretically, this requires the knowledge of unknown parameters such as the relative pose of the lighting between both the scene and the camera, the normal of the point being observed and the albedo of this point. However, it seems to be possible to extract local realistic estimations of these parameters from a sequence of images [38]. The knowledge of these parameters and their use in the interaction matrix should greatly increase the applications field of visual servoing.

6.1.3. Control of aircrafts by visual servoing

Participants: Odile Bourquardez, François Chaumette.

This study aims at developing visual servoing techniques for the control of the motions of aircrafts. As for fixed wing aircrafts, the considered applications are automatic landing, take off, and in-flight refueling. This year, we have concentrated our works on the two first tasks. Visual features based on the measurements that can be extracted from the image of the runway (typically, its border lines and start and end limits) have been modeled. A flight simulator provided by Dassault Aviation has been coupled to the visual servoing simulation software included in ViSP2, and classical control schemes have been implemented to track the trajectory of the runway in the image during take off and landing. Current works consist in building the kinematics and dynamics model of the aircraft, and in designing control schemes for such an underactuated system.

As for helicopters, we have designed and compared several control schemes based on the centroid of a target expressed in a spherical coordinate system, for positioning and stabilization tasks. The goal was to find a couple of visual feature and control scheme so that the sensitivity to any translational motion was the same. We have also implemented a 3D localisation algorithm on the X4-flyer developed at CEA-List within the Robvolint project (see Section 8.1.1). It uses a miniature camera, with wireless video transmission, mounted on the X4-flyer. Current works aim at testing if the data provided by the camera are sufficiently accurate, stable and not too much delayed to be used as input of the helicopter controller.

6.1.4. Vision-based tasks sequencing

Participants: Nicolas Mansard, François Chaumette.

The aim of this study is to realize global and high level robotics tasks by an adequate sequencing of local vision-based tasks. Classical visual servoing approaches tend to constrain all the degrees of freedom (DOF) of the robot during the execution of a task. The key idea is to control the robot with a very under-constrained task when it is far from the desired position, and to incrementally achieve the global task by adding further tasks in a stack as the robot moves closer to the goal. We have first used the classical redundancy formalism of the task function approach for such a purpose. As long as they are sufficient, the remaining DOF are used to take into account additional constraints, such as joint limits avoidance, preserving the visibility of the target, and avoiding occlusions. Closer from the goal, when not enough DOF remain available to consider these constraints, an execution controller selects a task to be temporary removed from the stack. The released DOF can then be used to ensure the additional constraints. A complete solution to implement this general idea has been implemented [27]. Experiments that prove the validity of the approach have also been performed.

This year, we have also revisited the classical redundancy formalism that allows to simultaneously consider a main task and additional constraints. Classically, the additional constraints are taken into account with the DOF which are not constrained by the main task. This solution may not be satisfactory if not enough DOF are available. We have thus proposed a new general method, called directional redundancy, that frees up some of the DOF constrained by the main task in addition of the remaining DOF. The general idea is to enable the motions produced to ensure the additional constraints that help the main task to be completed faster [24], [25]. In a formal framework, a non linear projection operator is built which ensures that the secondary control law does not disturb the main one. Experiments validating this approach have been realized: the visual servoing framework has been used to position a 6-DOF robot while simultaneously avoiding occlusions and joint limits.

Finally, through a collaboration with Manuel Lopes and José Santos-Victor from IST/ISR in Lisbon (see Section 8.2.2), we have applied the redundancy and task sequencing approaches on coarsely calibrated robots, using tasks jacobians estimated online. On such robots, we have shown that learning improves the servoing performance when task sequencing is used. Conversely, sequencing improves the convergence of jacobian learning, especially for tasks with several degrees of freedom. Eye-in-hand and eye-to-hand experiments have been realized with two robots with six degrees of freedom.

6.1.5. Qualitative visual servoing

Participants: Anthony Remazeilles, Nicolas Mansard, François Chaumette.

As mentioned in Section 3.1, classical visual servoing control laws are designed so that the current visual features $\mathbf{s}(t)$ reach a desired value $\mathbf{s}^*$. Nevertheless, obtaining exactly a desired value is a strong constraint that may be useless in some applications. For example, if an object tracking task is just specified as keeping the image of the object within the camera field of view, it is not necessary to require the center of gravity of this object to project exactly in the center of the image. Indeed, it is sufficient to ensure that the center of gravity belongs to an area around the center of the image.

For such a purpose, we have developed a new approach, named qualitative visual servoing, which ensures that the current feature reaches a confident interval defined in the vicinity of $\mathbf{s}^*$, that is such that $\mathbf{s}(t) \in [(1-p)\mathbf{s}^*, (1+p)\mathbf{s}^*]$ where scalar p defines the size of the interval. Such qualitative visual features can be combined with classical ones, if needed. As first experiments, it has been used to enforce points to stay within the camera field of view during classical image-based positioning tasks.

Our current works are devoted to the other issues that could benefit from this new approach. For example, joint limits and obstacles avoidances are examples of constraints that could be also considered. Furthermore, enlarging the desired visual features to intervals enables to release some robot degrees of freedom, which is of particular interest within the redundancy framework.

6.1.6. Navigation from a visual memory

Participants: Anthony Remazeilles, François Chaumette.

This research study is carried out in collaboration with P. Gros from Texmex project. Navigation is considered within a topological framework, in which the environment is described by an image data-base acquired during an off-line learning step. Within this context, the localization of the robotic system is nothing but the search in the data-base of images that are similar to the one describing the current position of the system. This is performed by using image retrieval techniques. The definition of the navigation path is then obtained by performing a search path into a graph associated to the image data base. This image path describes visually the environment the robot has to go through. Specific visual features have been defined to control the robot motion from the set of points in the current image that are matched with the consecutive images in the visual path. A control law derived from the qualitative visual servoing approach manages to make the robot progress along the path without being imposed to reach each position associated to the images of the path.

Results obtained on a robot arm for planar motion, and in simulation for $3D$ motions and $3D$ environments have confirmed the validity of our approach [32]. We are now considering the application of this formalism onto our real car-like robot, the Cycab. The different steps involved within the control loop are thus currently revisited in order to be robust to outdoor environments, and well-adapted to this kind of robot. The first step considered and resolved this year is the point tracking, which is detailed in Section 6.3.1.

6.2. Medical robotics

6.2.1. Automatic calibration of a robotized 3D ultrasound imaging system by visual servoing

Participants: Alexandre Krupa, François Chaumette.

Three-dimensional ultrasound imaging is used in numerous medical applications such as vascular imaging, cardiology, obstetrics, and neurosurgery. For example, this imaging modality is used to detect or survey the evolution of atherom plates attached to the walls of arteries that can be deadly for the patient. Two different techniques are used to perform 3D ultrasound imaging. The first one uses a 3D ultrasound sensor, but it currently provides low voxel resolution and a small field of view. The second technique, referred to as "3D free-hand ultrasound imaging", consists of measuring the relative displacement between each image captured by a two-dimensional ultrasound system in order to position them in a 3D frame. Usually the clinician performs the procedure manually through the use of an optical or magnetic localization system attached to the ultrasound probe. To assist the clinician, our purpose is to develop a robotic system which automatically moves the ultrasound probe during medical examinations in order to optimize the 3D ultrasound imaging. For example, this will allow for repeat examinations of a patient on different dates in order to observe quantitatively the pathology evolution under the same conditions.

As for manual 3D ultrasound imaging, it is crucial to know precisely the spatial calibration parameters of the ultrasound system in order to perform accurate 3D imaging. These parameters are usually identified during an initial step where the clinician has to capture a set of 2D ultrasound images of a known object for different positions and orientations of the probe. This procedure is laborious for the clinician and will influence the accuracy of the 3D ultrasound imaging. Therefore, we have proposed to automate the spatial calibration of a robotized 3D ultrasound system. A robotic task has been developed to automatically position the ultrasound probe on the intersection plane of a cross-wire phantom. To perform this task, a new visual servoing technique based on 2D ultrasound images was proposed in [21] to control automatically the motion of the ultrasound probe.

6.2.2. Ultrasound image-based visual servoing

Participant: Alexandre Krupa.

In the literature, very few researches deal with ultrasound image-based control of a medical robot holding the ultrasound probe. The issue of our study is to directly control the motion of the robot from visual features extracted from ultrasound images. The modeling aspects differ from the ones of classical visual servoing because the ultrasound transducer does not involve any perspective projection. In fact, an ultrasound probe provides information only in its observation plane whereas a standard camera provides a projection of the 3D scene to a 2D image. This year, we have been interested in modeling the coupling between the ultrasound probe and egg shaped objects. The choice of the egg shape is motivated by the likeness of this form to usual tumors. After the modeling step, we have developed a robotic task that consists in reaching a desired ultrasound image from an arbitrary initial position. The medical application may be to reach a tumor cross section having the maximum likelihood with a registered one. This will allow the radiologist to analyze quantitatively the tumor evolution using similar 2D ultrasound views captured on different dates. As simulation results of the visual servo control were promising, we now plan to develop a robust and real-time image processing algorithm to extract the visual features needed to validate experimentally this new approach with our medical robotic platform.

6.3. Visual tracking

Elaboration of objects tracking algorithms in image sequences is an important issue for research and application related to visual servoing and more generally for robot vision. A robust extraction and real-time spatio-temporal tracking process of visual cue is indeed one of the keys to success a visual servoing task. Fiducial markers have been used for a long time since they ensure a reliable and fast tracking. However, since such features are not present in realistic environments, it is no longer possible to be limited to such techniques.

6.3.1. 2D points tracking for mobile robot navigation

Participants: Sinisa Segvic, Anthony Remazeilles.

Landmark tracking is a critical issue in vision-based robot navigation. One needs image measurements that are robustly located, but also that have a sufficiently long life-time in order to define the robot motion during all the navigation.

The developed application considered the Harris points located at the pixels with high curvature in both dimensions of the signal associated to the grey-level image. A classical procedure for monitoring the tracking quality consists in requiring that the current measurement nicely warp towards its appearance in a reference frame. This procedure is valid when considering regions of interest that are projection of planar 3D patches. However, in real applications, a non negligible amount of points contain occluding boundaries. This fact becomes even more problematic when considering a camera on-boarded in car moving along straight sections of a road, since the motion of the camera induces a permanent increase of the apparent feature size. To deal with such points, a multi-scale monitoring has been implemented which maximizes the mesurements lifetime, while also detecting the tracking failures. A technique has also been proposed for inferring the parts of the reference which are not projected from the same 3D surface as the patch that has been consistently tracked until the present moment. Experiments realized on real sequences taken from cars driving through urban environments have shown that the proposed method is quite effective, especially in sequences with occlusions and large photometric variations.

6.3.2. 3D model-based tracking by virtual visual servoing

Participants: Andrew Comport, Cédric Riou, Éric Marchand, François Chaumette.

A real-time, robust and efficient 3D model-based tracking algorithm for a monocular vision system has been developed over the three last years [12]. Tracking objects in the scene amounts to compute the pose between the camera and the object. Non-linear pose computation is formulated by means of a virtual visual servoing and statistically robust estimation techniques have been introduced in the control law [15]. Articulated objects have also been considered [13]. This year we provided a theoretical comparison with other trackers [20] along with a quantitative evaluation of the tracker.

6.3.3. 2D model-free multimodal tracking

Participants: Muriel Pressigout, Éric Marchand.

For the time-being, most of the available tracking techniques can be divided into two main classes: edge-based and texture-based tracking. The former approach focuses on tracking 2D or 3D features such as geometrical primitives (points, segments, circles,...), object contours, 3D object, etc. The latter explicitly uses a texture or the luminance information that characterize the tracked object. These two classes of approaches have complementary advantages and drawbacks, so it is quite natural to use both of them. We have proposed in [29] [30] a reliable tracking for markerless planar objects based on the fusion of visual cues and on the estimation of a 2D transformation (an homography). The parameters of this transformation are estimated using a non-linear minimization of a unique criterion that integrates information on both texture and edges of the tracked object. Since data are likely to be corrupted with noise, we have considered a robust M-estimation into the minimization process through an iteratively re-weighted least squares implementation. Since this tracker is accurate and fast enough, it has been used successfully in visual servoing experiments whereas single cue trackers have failed or were not accurate enough.

Let us note that the considered 2D transformation is valid only for planar object. Since there does not exist any 2D transformation that accounts for a rigid motion of 3D object, we are also interested in multimodal 3D tracking, by combining pose computation and motion estimation, as described in the next section.

6.3.4. 3D model-based multimodal tracking

Participants: Muriel Pressigout, Éric Marchand.

Extending the approach presented in the previous paragraph, we have developed this year a real-time, robust and efficient 3D model-based tracking algorithm. A non linear minimization approach is used to register 2D and 3D cues for monocular 3D tracking. This approach fuses into a single non-linear objective function a 3D model-based approach based on the edge extraction (see paragraph 6.3.2), and a temporal matching relying on the texture analysis. Indeed, estimating both pose and camera displacement introduces an implicit spatio-temporal constraint the simple 3D model-based tracker lacks of. Furthermore, fusing measurements based on texture and edges improves the robustness of the tracking.

6.3.5. Object recognition

Participants: Claire Dune, Éric Marchand.

This new study is dedicated to object manipulation by visual servoing. The goal of this project, realized in cooperation with CEA/List (see Section 7.2), is to allow disable persons to grasp an object with the help of a robotic arm mounted on a wheel chair. This task should be achieved with a minimum of a priori information regarding the environment, the considered object, etc. We will start this study by looking for object recognition algorithms able to initalize automatically the tracking algorithms described in the previous sections.

7. Contracts and Grants with Industry

7.1. France Telecom R&D: Augmented reality in urban environment

Participant: Eric Marchand.

no. Inria 46134260, duration : 18 months.

This contract started in October 2005. Considering the algorithm described in 6.3.2, the goal of this project is to evaluate its capability to handle rough models provided by geographic information systems. A second part of this project will consider the automatic initialization of the tracking process.

7.2. CEA List: Clickrog: Object grasping for disable persons

Participants: Claire Dune, Eric Marchand.

duration : 36 months.

This contract started in November 2005. It is also supported by the Brittany Council through a grant to Claire Dune for her Ph.D. ("krog" means grasping in the Breton language). It is dedicated to object manipulation by visual servoing. The goal of this project is to allow disable persons to grasp an object with the help of a robotic arm mounted on a wheel chair. This task should be achieved with a minimum of a priori information regarding the environment, the considered object, etc.

7.3. European initiatives

7.3.1. ESA/Trasys: Vimanco: Vision-based manipulation of non-cooperative objects

Participants: Fabien Dionnet, Eric Marchand, François Chaumette.

duration : 12 months.

We began in September 2005 a project for the European Space Agency. It will be realized in collaboration with the Trasys company (Brussels), Galileo Avionica (Milano) and KUL (Leuven). Its aims is to develop a demonstrator of a robot arm in space environment able to grasp objects by visual servoing. The considered robot is the ESA Eurobot that should be on the International Space Station in 2008. Our task in this project is to provide algorithms for objects tracking and vision-based control.

7.4. French national programme Riam

7.4.1. Sora project

Participants: Cédric Riou, Andrew Comport, Éric Marchand.

no. Inria 2 03 C 1428, duration : 30 months.

The goal of the Sora project was to develop an object tracking algorithm for augmented reality applications. Our partners were Total-Immersion and VideoMage companies. Augmented reality has now progressed to the point where real-time applications are being considered and needed. On the other hand, it is important that synthetic elements are rendered and aligned in the scene in an accurate and visually acceptable way. In order to address these issues in real-time, a robust and efficient 3D model-based tracking algorithm has been proposed for a "video see through" monocular vision system. Virtual objects can then be projected into the scene from the camera pose computed using our algorithm. The tracking rate is 50Hz.

7.5. French national programme Predit

7.5.1. Mobivip project

Participants: Anthony Remazeilles, Sinisa Segvic, François Chaumette.

no. Inria 2 03 A 2005, duration : 36 months.

This project is a large project headed by Inria Sophia Antipolis. It is concerned with the navigation of mobile vehicles in urban environments. Within this project, our work consists in designing autonomous vision-based navigation techniques using an image database of the environment (see Sections 6.1.6 and 6.3.1). As for scientific aspects, this project is closely related to the Robea Bodega project described in Section 8.1.1.

8. Other Grants and Activities

8.1. National initiatives

8.1.1. Robea Robotics research programme

- *Parknav project: Complex dynamic scene interpretation and reactive motion planning.*
  Participant: Eric Marchand.
  This three-years project, started in October 2002, is a collaboration between Inria Rhône-Alpes (Sharp, Movi and Prima teams), the Laas (Ria group) and our group. The objective of this project is the automatic control of vehicles in parking environment using outside cameras. For this year, we have tested the image motion detection and estimation algorithms on image sequences acquired on the platform available in Grenoble.
- *Egocentre project*
  Participants: Nicolas Mansard, François Chaumette.
  This three-years project, started in September 2003, aims at developing task sequencing schemes to realize high level robotics tasks from local sensor-based control techniques. It involves a collaboration between Laas and Cerco in Toulouse, Enit in Tarbes, and our team. Nicolas Mansard's PhD is realized in the scope of this project (see Section 6.1.4).

- *Bodega project*
  Participants: Anthony Remazeilles, Sinisa Segvic, François Chaumette.
  This two-years project, started in November 2003, aims at designing vision-based and sensor-based methods for the autonomous navigation of mobile vehicles moving around an urban environment. It involves a collaboration between Ensil in Limoges, UTC in Compiègne, Lasmea in Clermont-Ferrand, the Icare team of Inria Sophia-Antipolis, and our team. Works described in Sections 6.1.6 and 6.3.1 are achieved within this project.
- *Robvolint project*
  Participants: Odile Bourquardez, François Chaumette.
  This two-years project, started in November 2003, aims at developing vision-based localization techniques and visual servoing schemes for small helicopters moving around an indoor environment. It involves a collaboration between I3S in Nice, CEA in Fontenay-aux-Roses, Ircynn in Nantes, and our team. Works described in Section 6.1.3 about helicopters are realized within this project.

8.2. Bilateral international co-operation

8.2.1. CNRS PICS, France-Australia

Participants: Odile Bourquardez, François Chaumette.

This international collaboration between France and Australia is supported by CNRS. It is about visual servo-control of unmanned aerial vehicles. It started fall 2005 for three years. It joins Rob Mahony (Australian National University, Canbera), Peter Corke and Jonathan Roberts (CSIRO, Melbourne), Tarek Hamel (I3S, Sophia-Antipolis), Vincent Moreau (CEA-List, Paris) and our group.

8.2.2. Inria-Grices project, France-Portugal

Participants: Nicolas Mansard, François Chaumette.

This collaboration with IST Lisbon, Portugal (Prof. J. Santos-Victor) is concerned with visual servoing for robotics applications. M. Lopes did a one month visit in our group in February 2005, and N. Mansard did a one month visit at IST in June 2005. The works realized during these visits are described at the end of Section 6.1.4.

8.2.3. Visiting scientists

- Jorge Pomares, Assistant Professor at University of Alicante, Spain, did a one-month visit in March 2005.
- Short visits by Gabriel Lopes (University of Michigan), Pascal Fua (EPFL Lausanne), Brad Nelson (ETHZ Zurich), Daisuke Kotake (Canon Inc., Tokyo), Kui Yan (Chinese Academy of Sciences, Beijing).

9. Dissemination

9.1. Leadership within scientific community

- F. Chaumette is a member of the CNRS Experts Committee on Robotics, in charge of the prospective in this field.
- E. Marchand is a member of the Scientific Council of the University of Rennes 1.
- F. Spindler is a member of the engineers reviewing committee of the French institute for agronomy research (Inra).
- F. Chaumette is a member of the Specialist Committee of IFSIC.

- *Editorial boards of journals*
  - F. Chaumette has been Associate Editor of IEEE Trans. on Robotics until July 2005.
- *Conference organization*
  - F. Chaumette and D. Duhaut (UBS Lorient) have been the General Chairs of JNRR'05 ("Journées Nationales de la Recherche en Robotique") which have been held in October 2005 in Guidel (Morbihan) [11]. More than 160 people from 50 french labs have participated to this conference. Stéphanie Lemaile, Eric Marchand, Fabien Spindler, Muriel Pressigout and Nicolas Mansard were in the Local Organizing Committee.
- *Technical program committees of conferences*
  - F. Chaumette : ICRA'2005, CVPR'2005, ICCV'2005, ICRA'2006, ECCV'2006, CVPR'2006, IROS'2006.
  - E. Marchand : ORASIS'2005, ISMAR'2005, ECCV'2006.

9.2. Teaching

- Master M2RI of Computer Science, Ifsic, University of Rennes 1 (E. Marchand): 3D Computer vision.
- Diic INC, Ifsic, University of Rennes 1 (E. Marchand, F. Chaumette: 3D vision, visual servoing; E. Marchand, F. Spindler: programming tools for image processing).
- Insa Rennes, Electrical Engineering Dpt (F. Chaumette, F. Spindler, O. Bourquardez: computer vision).
- International Online Course on Visual Servoing Techniques (F. Chaumette)
- Graduate student interns: W. Bachta (Master LSIIT, Strasbourg), M. Collouard (Master STIR Rennes), L. Parada (Supélec Rennes), C. Renaudin (Insa Rennes)
- External thesis supervision:
  - J. Pages (University of Girona, Spain and Cemagref Rennes) co-supervised by F. Chaumette.

9.3. Participation in seminars, invitations, awards

- Paper [21] has been selected as one of the ten finalist papers for the Best Paper Award at IROS'2005 (award will be nominated in 2006).
- Paper [33] has been selected as one of the three finalist papers for the Best Vision Paper Award at ICRA'2005.
- F. Chaumette was invited in April 2005 at IST, Lisbon, as external reviewer of Paulo Ferreira and Paulo Goncalves's Ph-D thesis defense,
- F. Chaumette has been invited for a three-day visit at the Beckmann Institute in July 2005.
- Eric Marchand has presented a demonstration of the augmented reality Markerless software at Siggraph conference in Los Angeles in August 2005, and at the 14th Science Fair at the Luxembourg Gardens in Paris in October 2005.

10. Bibliography

Major publications by the team in recent years

[1] F. CHAUMETTE. *Image moments: a general and useful set of features for visual servoing*, in "IEEE Trans. on Robotics and Automation", vol. 20, n^o 4, August 2004, p. 713-723.

[2] F. CHAUMETTE. *Potential problems of stability and convergence in image-based and position-based visual servoing*, in "The Confluence of Vision and Control, LNCIS Series, No 237, Springer-Verlag", 1998, p. 66-78.

[3] A.I. COMPORT, E. MARCHAND, F. CHAUMETTE. *A real-time tracker for markerless augmented reality*, in "ACM/IEEE Int. Symp. on Mixed and Augmented Reality, ISMAR'03, Tokyo, Japan", October 2003, p. 36–45.

[4] A. CRÉTUAL, F. CHAUMETTE. *Visual servoing based on image motion*, in "Int. Journal of Robotics Research", vol. 20, n^o 11, November 2001, p. 857–877.

[5] B. ESPIAU, F. CHAUMETTE, P. RIVES. *A new approach to visual servoing in robotics*, in "IEEE Trans. on Robotics and Automation", vol. 8, n^o 6, June 1992, p. 313-326.

[6] E. MALIS, F. CHAUMETTE, S. BOUDET. *2 1/2 D visual servoing*, in "IEEE Trans. on Robotics and Automation", vol. 15, n^o 2, April 1999, p. 238-250.

[7] E. MARCHAND, F. CHAUMETTE. *Virtual Visual Servoing: a framework for real-time augmented reality*, in "Eurographics 2002 Conf. Proc., Saarebrücken, Germany", Computer Graphics Forum, vol. 21, n^o 3, 2002, p. 289–298.

[8] E. MARCHAND, F. CHAUMETTE. *An autonomous active vision system for complete and accurate 3D scene reconstruction*, in "Int. Journal of Computer Vision", vol. 32, n^o 3, August 1999, p. 171-194.

[9] E. MARCHAND, N. COURTY. *Controlling a camera in a virtual environment*, in "The Visual Computer Journal", vol. 18, n^o 1, February 2002, p. 1-19.

[10] Y. MEZOUAR, F. CHAUMETTE. *Path planning for robust image-based control*, in "IEEE Trans. on Robotics and Automation", vol. 18, n^o 4, August 2002, p. 534-549.

Books and Monographs

[11] F. CHAUMETTE, D. DUHAUT (EDS.). *5e Journées Nationales de la Recherche en Robotique*, IRISA, Université de Rennes 1, 315 p., Guidel, Morbihan, October 2005, http://jnrr05.irisa.fr.

Doctoral dissertations and Habilitation theses

[12] A.I. COMPORT. *Robust real-time 3D tracking of rigid and articulated objects for augmented reality and robotics*, Ph. D. Thesis, Université de Rennes 1, Mention informatique, September 2005.

Articles in referred journals and book chapters

[13] A.I. COMPORT, E. MARCHAND, F. CHAUMETTE. *Complex articulated object tracking.*, in "Electronic Letters on Computer Vision and Image Analysis, ELCVIA, Special Issue on "Articuled Motion & Deformable Objects"", vol. 5, n° 3, May 2005, p. 20-30.

[14] A.I. COMPORT, E. MARCHAND, F. CHAUMETTE. *Efficient model-based tracking for robot vision.*, in "Advanced Robotics, Special Issue on "Selected papers from IROS'04"", vol. 16, n° 10, December 2005, p. 1097-1113.

[15] A.I. COMPORT, E. MARCHAND, F. CHAUMETTE. *Statistically robust 2D visual servoing*, in "IEEE Trans. on Robotics", to appear, 2006.

[16] H. HADJ ABDELKADER, Y. MEZOUAR, P. MARTINET, F. CHAUMETTE. *Asservissement visuel en vision omnidirectionnelle à partir de droites*, in "Traitement du Signal, numéro spécial sur la vision omnidirectionnelle", vol. 22, n° 5, September 2005.

[17] E. MARCHAND, F. CHAUMETTE. *Feature tracking for visual servoing purposes*, in "Robotics and Autonomous Systems, Special Issue on “Advances in Robot Vision”", vol. 52, n° 1, July 2005, p. 53-70.

[18] E. MARCHAND, F. SPINDLER, F. CHAUMETTE. *ViSP: a generic software platform for visual servoing*, in "IEEE Robotics and Automation Magazine, Special Issue on "Software Packages for Vision-Based Control of Motion"", vol. 12, n° 4, December 2005.

[19] O. TAHRI, F. CHAUMETTE. *Point-based and Region-based Image Moments for Visual Servoing of Planar Objects*, in "IEEE Transactions on Robotics", vol. 21, n° 6, December 2005, p. 1116-1127.

Publications in Conferences and Workshops

[20] A.I. COMPORT, D. KRAGIC, E. MARCHAND, F. CHAUMETTE. *Robust Real-Time Visual Tracking: Comparison, Theoretical Analysis and Performance Evaluation*, in "IEEE Int. Conf. on Robotics and Automation, ICRA'05, Barcelona, Spain", April 2005, p. 2852-2857.

[21] A. KRUPA, F. CHAUMETTE. *Control of an Ultrasound Probe by Adaptive Visual Servoing*, in "IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'05, Edmonton, Canada", vol. 2, August 2005, p. 2007-2012.

[22] P. LI, O. TAHRI, F. CHAUMETTE. *A shape tracking algorithm for visual servoing.*, in "IEEE Int. Conf. on Robotics and Automation, ICRA'05, Barcelona, Spain", April 2005, p. 2858-2863.

[23] E. MALIS, E. MARCHAND. *Méthodes robustes d'estimation pour la vision robotique*, in "Journées nationales de la recherche en robotique, JNRR'05, Guidel, France", October 2005, p. 51-60.

[24] N. MANSARD, F. CHAUMETTE. *A new redundancy formalism for avoidance in visual servoing*, in "IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'05, Edmonton, Canada", vol. 2, August 2005, p. 1694-1700.

[25] N. MANSARD, F. CHAUMETTE. *Directional redundancy: a new approach of the redundancy formalism*, in "IEEE Conf. on Decision and Control and European Control Conference, CDC/ECC 2005, Seville, Spain", December 2005, p. 5366-5371.

[26] N. MANSARD, F. CHAUMETTE. *Enchaînement de tâches en asservissement visuel : application à l'évitement des butées articulaires*, in "Congrès jeunes chercheurs en Vision par ordinateur, ORASIS'2005, Clermont-Ferrand, France", May 2005.

[27] N. MANSARD, F. CHAUMETTE. *Visual servoing sequencing able to avoid obstacles.*, in "IEEE Int. Conf. on Robotics and Automation, ICRA'05, Barcelona, Spain", April 2005, p. 3154-3159.

[28] J. PAGÈS, C. COLLEWET, F. CHAUMETTE, J. SALVI. *Robust Decoupled Visual Servoing Based on Structured Light*, in "IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'05, Edmonton, Canada", vol. 2, August 2005, p. 2676-2681.

[29] M. PRESSIGOUT, E. MARCHAND. *A model free hybrid algorithm for real time tracking*, in "IEEE Int. Conf. on Image Processing, ICIP'05, Genoa, Italia", vol. 3, September 2005, p. 97-100.

[30] M. PRESSIGOUT, E. MARCHAND. *Real-time planar structure tracking for visual servoing: a contour and texture approach.*, in "IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'05, Edmonton, Canada", vol. 2, August 2005, p. 1701-1706.

[31] M. PRESSIGOUT, E. MARCHAND. *Suivi temps-réel d'objet plan: approche hybride contour/texture*, in "Congrès jeunes chercheurs en Vision par ordinateur, ORASIS'2005, Fournol, France", May 2005.

[32] A. REMAZEILLES, F. CHAUMETTE, P. GROS. *Image based robot navigation in 3D environments*, in "Int. Symp. on Optomechatronic Technologies, ISOT'05, Sapporo, Japan", December 2005.

[33] O. TAHRI, F. CHAUMETTE. *Complex objects pose estimation based on image moment invariants.*, in "IEEE Int. Conf. on Robotics and Automation, ICRA'05, Barcelona, Spain", April 2005, p. 438-443.

Internal Reports

[34] J. PAGÈS, C. COLLEWET, F. CHAUMETTE, J. SALVI. *Visual servoing by means of structured light for plane-to-plane positioning*, Technical report, n° 5579, INRIA, May 2005.

[35] M. PRESSIGOUT, E. MARCHAND. *Real Time Planar Structure Tracking: a Contour and Texture Approach*, Technical report, n° 1698, IRISA, March 2005.

Miscellaneous

[36] F. SPINDLER, C. BLONZ. *Visual Servoing Research Activities in Lagadic, JNRR'05 Video session, Guidel, Morbihan*, October 2005.

Bibliography in notes

[37] S. HUTCHINSON, G. HAGER, P. CORKE. *A tutorial on visual servo control*, in "IEEE Trans. on Robotics and Automation", vol. 12, nº 5, October 1996, p. 651-670.

[38] S. NEGAHDARIPOUR. *Revised definition of optical flow: integration of radiometric and geometric cues for dynamic scene analysis*, in "IEEE Trans. on Pattern Analysis and Machine Intelligence", vol. 20, nº 9, September 1998, p. 961–979.

[39] C. SAMSON, M. LE BORGNE, B. ESPIAU. *Robot Control: the Task Function Approach*, Clarendon Press, Oxford, 1991.