Vision/Force Control in Task-Oriented Grasping and Manipulation

Mario Prats[1,3], Philippe Martinet[1,2], Angel P. del Pobil[3] and Sukhan Lee[1]

***Abstract*— In this paper, we present a novel approach for sensor-guided robotic execution of everyday tasks, which is amenable to be integrated in current mobile manipulators and humanoid robots. We consider a robot which is observing simultaneously his hand and the object to manipulate, by using an external camera (i.e. robot head). Task-oriented grasping algorithms are used in order to plan a suitable grasp on the object according to the task to perform. A new vision/force coupling approach [1] is used in order to, first, guide the robot hand towards the grasp position and, second, perform the task taking into account external forces. Experimental results on a real robot are presented which validate our approach.**

I. INTRODUCTION

There is a need for fully autonomous robots that can perform a great variety of tasks, as humans do. In the literature, many different mobile manipulators have been presented, aimed to combine manipulation capabilities with mobility. In the last years, humanoid robots have become very popular in the robotics community, also trying to combine manipulation and mobility, but aimed to coexist with humans in the future.

For humanoid robots, advanced manipulation skills are very important. Either if we want them to assist the elderly people at home, or to act as receptionists or workers, robust task-oriented grasping and manipulation algorithms must be developped and integrated into real systems.

In this paper, we present a novel approach for sensor-guided robotic task execution that is amenable to be integrated in current mobile manipulators and humanoid robots. We consider a robot which is observing simultaneously his hand and the object to manipulate, by using an external camera (i.e. robot head, see Figure 1). Already developped task-oriented grasping algorithms [2] are used in order to plan a suitable grasp on the object according to the task to perform. A new vision/force coupling approach [1] is used in order to, first, guide the robot hand towards the grasp position and, second, perform the task taking into account external forces.

The problem of hand/object alignement for grasping tasks has been addressed by other authors. In [3], the authors presented a visual servoing framework for aligning the end-effector with an object. Instead of working in the euclidean space, visual servoing was done on the projective space by doing projective reconstruction with a stereo camera, thus avoiding the need for camera calibration. The desired gripper-to-object relationship was learnt during an off-line procedure. In [4], an external position-based visual servoing approach was used on a humanoid robot in order to guide the hand towards the object. Hand pose was estimated by a kalman filter taking as input the stereo reconstruction of a set of LEDs attached on the robot hand.

[1]Intelligent System Research Center, Sungkyunkwan University, Suwon, Korea.
[2]LASMEA, Blaise Pascal University, Clermont-Ferrand, France.
[3]Robotic Intelligence Lab, Jaume I University, Castellón, Spain.

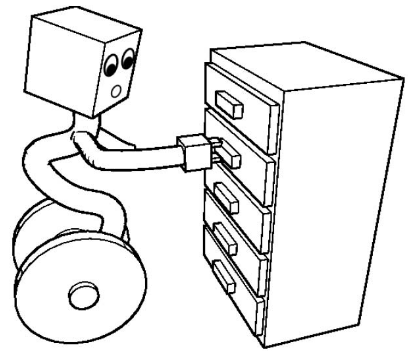

Fig. 1. Considered scenario.

As in [4], we also adopt a position-based visual servoing control law, because of the facilities that this approach offers for task specification. Instead of using a stereo camera and perform 3D reconstruction, we make use of a single camera and follow the virtual visual servoing approach for pose estimation [5]. The goal of the vision control loop is to align the gripper with respect to some part of the object (i.e, handle). As the pose of the gripper and the object is estimated online, the relative position between both can be computed at each iteration without the need of knowing the position of the camera with respect to the robot base. Therefore, the robot is still able to perform the task even in the presence of low camera motion. Task execution is independent of camera position. No extrinsic camera parameters are needed, which makes the integration of this approach into existing real robotic systems very easy, and opens the door to best-view planning algorithms for head control. In addition, instead of learning the grasp position during an offline stage like in [3], we make use of a task-oriented grasp planning algorithm [2] which autonomously computes which part of the object should be grasped in order to perform a given task. Finally, and in contrast with existing works, our visual servoing does not finish when the robot grasps the object. Instead, a novel vision/force control framework is adopted in order to perform a given task on the object. Thus, visual servoing is not only used for hand-object alignment (reaching), but also for task execution and supervision (interaction).

Some researchers have addressed the problem of vision/force control and two main approaches (impedance-based and hybrid-based strategies) have been studied [6],

[7]. In these schemes the idea is merely to replace the classical position controller by a vision-based controller. Hybrid control separates vision control and force control into two separate control loops, that operate in orthogonal directions. With this approach, it is not possible to control a direction simultaneously in vision and force. With the impedance-based control, the six degrees of freedom can be simultaneously vision- and force-controlled. However, coupling is done at the control level and local minima can appear during convergence. Our approach for vision/force coupling [1], based on the concept of external control [8], does the coupling in sensor-space, which allows to control vision and force on all the degrees of freedom, whereas only the vision control law is directly connected to the robot.

In Section II, the general framework of our work is presented. Section III describes the theoretical basis of the vision/force coupling scheme that we propose. In Section IV, the theoretical framework is applied to a real task that is executed by a real robot, and experimental results are presented. Finally, some conclusions and future lines are outlined in Section V.

II. GENERAL FRAMEWORK

We consider the general case of a mobile manipulator (or humanoid) working in a home environment. We assume that the robot is endowed with an object recognition module, so that it is able to recognize the object to manipulate and to retrieve its geometrical and structural model from a database. The robot is able to move in front of the object by using navigation capabilities such as mapping, localization, obstacle avoidance, etc.

For the experimental validation, we have used a mobile manipulator composed of an Amtec 7DOF ultra light weight robot arm mounted on an ActivMedia PowerBot mobile robot. The hand of the robot is a PowerCube parallel jaw gripper. This robot belongs to the Intelligent Systems Research Center (Sungkyunkwan University, South Korea), and is already endowed with recognition and navigation capabilities [9].

A. Representation of objects

We are interested in the robotic manipulation of articulated objects that can be commonly found in our everyday life, such as doors, drawers, wardrobes, etc. For providing the robot with such advanced manipulation capabilities, we use a special kind of object models [2], that include:

- Geometrical information, used for object recognition purposes
- Kinematic information, or a description of the object mechanism, used for manipulation purposes.

We describe an object as a set of different parts that are assembled together. Each part is defined on its own reference frame, which is independent from the other parts. A set of relations is defined between the parts, in terms of constrained and free degrees of freedom, i.e. a motion constraint is defined with each frame.

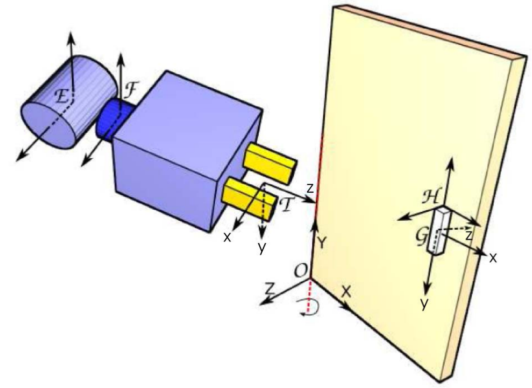

Fig. 2. Considered frames.

Figure 2 shows an example of a door representation. It is composed of two parts: the door table, defined in frame $\mathcal{O}$ -which is also the object reference frame- and the handle, defined in frame $\mathcal{H}$. The model, as described in [2], includes the relation between the different object parts. In this case, the relation between the handle and the door table is known, and represented as an homogeneous transformation matrix ${}^{\mathcal{O}}\mathbf{T}_{\mathcal{H}}$. The model also includes the degrees of freedom (motion constraint) for each part. In the example of Figure 2, the frame $\mathcal{H}$ is fixed with respect to $\mathcal{O}$, but the frame $\mathcal{O}$ has one degree of freedom: a rotation around Y axis, which corresponds to the task of opening the door. Thus, the task is specified to the robot by means of a frame and the degree of freedom that must be activated on it. For more details on the object representation, refer to [2].

B. Task-oriented grasp planning

Task-oriented grasp planning deals with the problem of finding a grasp on an object which is suitable for a particular task. There are few works about grasping that take the task into account [10], [11]. Most of them do not consider the task during grasp planning. Instead, the task is considered on the grasp evaluation stage as a quality measure. In practice, lots of grasps would have to be generated and evaluated, making these approaches computationally unaffordable. In [2], we presented a task-oriented grasp planning algorithm based on hand preshapes [12].

The input to this algorithm is the object model (as described previously) and the task to perform, in terms of a mechanism (i.e. degree of freedom) to be activated on the object. The algorithm provides the following:

- A grasp frame, $\mathcal{G}$ in the example of Figure 2, where the hand has to be moved, which is related with the object reference frame, $\mathcal{O}$, by computing the homogeneous transformation matrix ${}^{\mathcal{O}}\mathbf{T}_{\mathcal{G}}$.
- A hand preshape suitable for the task, including a tool frame, $\mathcal{T}$ in Figure 2, attached to the hand, which determines the control strategy that will be followed for grasping [2]. The tool frame is related to the end-effector frame by the transformation ${}^{\mathcal{E}}\mathbf{T}_{\mathcal{T}}$

For the particular case of a parallel jaw gripper, we only consider the precission hand preshape [2]. The tool frame, $\mathcal{T}$, is set to the middle point between both fingertips as shown in Figure 2.

III. VISION/FORCE CONTROL FOR DAILY TASKS

The task execution process for articulated objects can be divided into two stages:

- A reaching phase, where the hand of the robot must be moved towards the handle until the grasp is executed successfully.
- An interaction phase, where the hand is in contact with the object and a particular mechanism must be activated.

The reaching task can be performed by moving the tool frame towards the grasp frame. For this, we need to estimate the transformation ${}^{\mathcal{T}}\mathbf{T}_{\mathcal{G}}$ between both frames (i.e, the hand and the handle). One possibility is to estimate this transformation by using the robot localization algorithms, and to reach the handle in open-loop. But robot localization may contain important errors, and in practice this approach would fail. We propose a position-based visual servoing closed-loop approach where a robot head observes both the gripper and the object and tries to achieve a relative position between both.

Regarding the interaction phase, it is worth noting that, during this phase, the robot is in contact with the environment and some degrees of freedom are constrained. Errors in the models and in the vision control law would cause some small hand displacements along the constrained directions, which could generate very high external forces. It is necessary to take these forces into account and to locally modify the trajectory in order to minimize external forces at the same time that the vision control law guarantees the whole task execution.

In the following sections, the theoretical framework for our position-based visual servoing approach and vision/force control will be presented.

A. *External position-based visual servoing*

Several vision-based control laws have been proposed in the literature. They are generally classified in three groups, namely position-based, image-based and hybrid-based control. The first one works in 3D cartesian space and requires, in most cases, a model of the object and the camera intrinsic parameters [13]. In contrast, image-based visual servoing works directly in the image space [14]. More recently, several researchers have explored hybrid approaches which combine euclidean and image information [15].

As already mentionned in [4], the natural space for specifying the task is the cartesian space, and there are evidences that humans use 3D information for task planning. Thus, we adopt a position-based visual servoing approach using an external camera which observes simultaneously the gripper and the object. Note that this is the common configuration in humanoid robots.

We set the vector $\mathbf{s}$ of visual features to:

$$\mathbf{s} = \begin{pmatrix} \mathbf{t} \\ \mathbf{u}\theta \end{pmatrix}$$

where $\mathbf{t}$ is the translational part of the homogeneous matrix ${}^{\mathcal{T}}\mathbf{T}_{\mathcal{G}}$, and $\mathbf{u}\theta$ is the axis/angle representation of the rotational part of ${}^{\mathcal{T}}\mathbf{T}_{\mathcal{G}}$.

The matrix ${}^{\mathcal{T}}\mathbf{T}_{\mathcal{G}}$, which relates hand and handle, is computed directely from the visual observation of the gripper and object, according to the following expression:

$$\left({}^{\mathcal{C}}\mathbf{T}_{\mathcal{GP}} \cdot {}^{\mathcal{E}}\mathbf{T}_{\mathcal{GP}}{}^{-1} \cdot {}^{\mathcal{E}}\mathbf{T}_{\mathcal{T}}\right)^{-1} \cdot {}^{\mathcal{C}}\mathbf{T}_{\mathcal{OP}} \cdot {}^{\mathcal{O}}\mathbf{T}_{\mathcal{OP}}{}^{-1} \cdot {}^{\mathcal{O}}\mathbf{T}_{\mathcal{G}} \quad (1)$$

where ${}^{\mathcal{C}}\mathbf{T}_{\mathcal{GP}}$ is an estimation of the pose of an arbitrary hand frame, expressed in camera frame. ${}^{\mathcal{C}}\mathbf{T}_{\mathcal{OP}}$ is an estimation of an arbitrary object frame pose, expressed in camera frame. We are currently estimating hand and object pose by virtual visual servoing [5], using a set of point features drawn on a pattern whose model is known. One pattern is attached to the gripper, in a known position ${}^{\mathcal{E}}\mathbf{T}_{\mathcal{GP}}$. Another pattern is attached to the object, also in a known position with respect to the object reference frame: ${}^{\mathcal{O}}\mathbf{T}_{\mathcal{OP}}$. As future lines we would like to implement a feature extraction algorithm in order to use natural features of the object instead of the markers. Note that this will not significantly affect the current implementation, as virtual visual servoing pose estimation can deal with different types of visual features [5]. Finally, the tool frame ${}^{\mathcal{E}}\mathbf{T}_{\mathcal{T}}$ and the grasp frame ${}^{\mathcal{O}}\mathbf{T}_{\mathcal{G}}$ are computed by the task-oriented grasp planning algorithm presented in section II-B and detailed in [2].

We compute the velocity in the tool frame $\tau_{\mathcal{T}}$ using a classical visual servoing control law:

$$\tau_{\mathcal{T}} = -\lambda \mathbf{e} + \widehat{\frac{\partial \mathbf{e}}{\partial t}} \quad (2)$$

where $\mathbf{e}(\mathbf{s}, \mathbf{s}^*) = \widehat{\mathbf{L}_{\mathbf{s}}^{+}}(\mathbf{s} - \mathbf{s}^*)$ (in our case, $\mathbf{s}^* = 0$). The last term, $\widehat{\frac{\partial \mathbf{e}}{\partial t}}$, is the estimation of how the visual features change over time. It is related to the object motion, and should be taken into account when the hand is in contact with the environment in order to reduce tracking errors. However, we can neglect it, because the use of force feedback allows us to cope with these small tracking errors. The interaction matrix $\widehat{\mathbf{L}_{\mathbf{s}}}$ is set for the particular case of position-based visual servoing:

$$\widehat{\mathbf{L}_{\mathbf{s}}} = \begin{pmatrix} -\mathbf{I}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & -\mathbf{L}_w \end{pmatrix}$$

$$\mathbf{L}_w = \mathbf{I}_{3\times3} - \frac{\theta}{2}[\mathbf{u}]_\times + \left(1 - \frac{\text{sinc}(\theta)}{\text{sinc}^2(\frac{\theta}{2})}\right)[\mathbf{u}]_\times^2$$

where $[\mathbf{u}]_\times$ is the skew anti-symmetric matrix for the rotation axis $\mathbf{u}$. Finally, the joint velocities that are sent to the robot are computed as:

$$\dot{\mathbf{q}} = \mathbf{J}^{-1} \cdot \mathbf{L}_\times \cdot \begin{pmatrix} {}^{\mathcal{E}}\mathbf{R}_{\mathcal{T}} & [{}^{\mathcal{E}}\mathbf{t}_{\mathcal{T}}]_\times \cdot {}^{\mathcal{E}}\mathbf{R}_{\mathcal{T}} \\ \mathbf{0}_{3\times3} & {}^{\mathcal{E}}\mathbf{R}_{\mathcal{T}} \end{pmatrix} \cdot \tau_{\mathcal{T}}$$

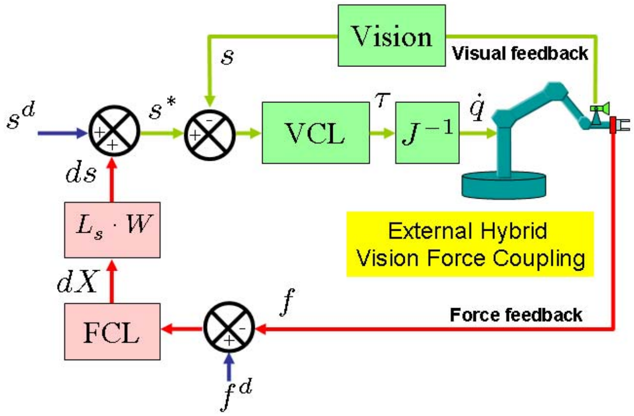

Fig. 3. External hybrid vision/force coupling.

where $\mathbf{J}$ is the robot jacobian and $\mathbf{L}_\times$ relates $\tau_\mathcal{E}$ and $\dot{\mathbf{X}}_\mathcal{E}$ according to $\dot{\mathbf{X}}_\mathcal{E} = \mathbf{L}_\times \cdot \tau_\mathcal{E}$ [13]. It is worth noting that, for very small displacements, $\mathbf{L}_\times$ can be taken as the identity matrix, and, thus, $\dot{\mathbf{X}}_\mathcal{E} = \tau_\mathcal{E}$. Finally, ${}^\mathcal{E}\mathbf{R}_\mathcal{T}$ and ${}^\mathcal{E}\mathbf{t}_\mathcal{T}$ are, respectively, the rotational and translational part of the homogeneous transformation matrix ${}^\mathcal{E}\mathbf{T}_\mathcal{T}$.

B. Vision/force control law

In [1], a new vision/force coupling approach is presented, where the force control loop is closed around an internal vision control loop in a hierarchical way (see Figure 3). The reference trajectory $\mathbf{s}^d$ used as original input of the vision-based controller is modified according to the external force control loop. The force control is performed by direct control: when the robot is moving from $\mathrm{d}\mathbf{X}$ against a contact surface, the force measurement is proportional to the environment stiffness $\mathbf{K}$ and the displacement $\mathrm{d}\mathbf{X}$. The desired vector of visual features $\mathbf{s}^d$ is then modified according to the force controller output, which is a relative position $\mathrm{d}\mathbf{X}$, projected on sensor space by means of the interaction matrix $\widehat{\mathbf{L}_\mathbf{s}}$. When the end-effector is not in contact with the external environment, the output of the force controller is null and the robot is controlled according to the vision-based controller output. Otherwise the force controller only modifies the reference trajectory of visual observations.

In the control scheme, shown in Figure 3, the desired wrench $\mathbf{f}^d$ is added as input in the force feedback control loop. The stiffness is controlled by the force controller (FCL) according to a proportional control law:

$$d\mathbf{X} = \mathbf{K}^{-1}(\mathbf{f}^d - \mathbf{f})$$

The force controller only modifies the reference trajectory of visual observations $\mathbf{s}^d$:

$$\mathbf{s}^* = \mathbf{s}^d + d\mathbf{s} \tag{3}$$

where $\mathbf{s}^*$ is the modified reference for visual features and $\mathrm{d}\mathbf{s}$ can be computed by projecting $\mathrm{d}\mathbf{X}$ by means of the interaction matrix as $d\mathbf{s} = \widehat{\mathbf{L}_\mathbf{s}} \cdot \mathbf{L}_\times^{-1} \cdot d\mathbf{X}$. It is worth noting that $d\mathbf{X}$ must be first transformed, from the force sensor frame, to the camera frame, via the corresponding screw transformation matrix.

The hierarchical juxtaposition of the force control loop on the vision control loop provides several advantages according to the existing methods [6], [7]: selection matrices and time-dependent geometric transformations are eliminated from the control loop leading to a controller design independent of the arm configuration. Since the force control only acts on the reference trajectory, conflicts between force and vision controllers are avoided [1].

IV. TASK EXECUTION

A. Task description

The vision/force control law of section III-B has been implemented on the robot and validated with the execution of a common manipulation task in our daily life: opening the door of a wardrobe. The model of the wardrobe is shown in Figure 2. The task to perform can be described in terms of the structural model as applying a rotational velocity around Y axis of frame $\mathcal{O}$.

The task-oriented grasp planning algorithm [2] computes the grasp frame $\mathcal{G}$ and the tool frame $\mathcal{T}$ and relates them to the object reference frame with ${}^\mathcal{O}\mathbf{T}_\mathcal{G}$, and to the end-effector frame with ${}^\mathcal{E}\mathbf{T}_\mathcal{T}$, respectively (see Figure 2).

B. Reaching

The reaching task is divided into three different subtasks: reaching a pre-grasp position, reaching the grasp position and performing the grasp. The robot switches from one subtask to another when the resulting velocity of the vision/force controller is close to zero (i.e, the desired references have been reached).

1) Reaching a pre-grasp position: A pre-grasp frame is computed by the task-oriented grasp planning algorithm [2], and it is related to the object reference frame with the transformation ${}^\mathcal{O}\mathbf{T}_\mathcal{P}$. The pre-grasp position is used in order to adopt an initial configuration with respect to the final grasp frame so that the robot can reach the handle from a good direction.

The transformation between the tool frame and the pre-grasp frame ${}^\mathcal{T}\mathbf{T}_\mathcal{P}$ is computed at each iteration by equation 1, and then used for building the visual features vector $\mathbf{s}$. During this step there is no contact with the environment. Thus, the force loop in the vision/force control law is not modifying the visual reference. This means that the system behaves according to the vision control law of section III-A

Figure 5 shows the evolution of the visual velocity for each degree of freedom. Initially, the tool frame is far from the pre-grasp frame (see Figure 4a), so that there is a large visual error. The visual control law makes this error converge to zero, which corresponds to the situation where the tool frame matches with the pre-grasp frame (see Figure 4b). The complete sequence can be seen in the video accompanying this paper.

Fig. 4. Reaching the handle. (a) Initial position (b) Reaching the pre-grasp position (c) Reaching the grasp position (d) Grasping (e) Interaction.

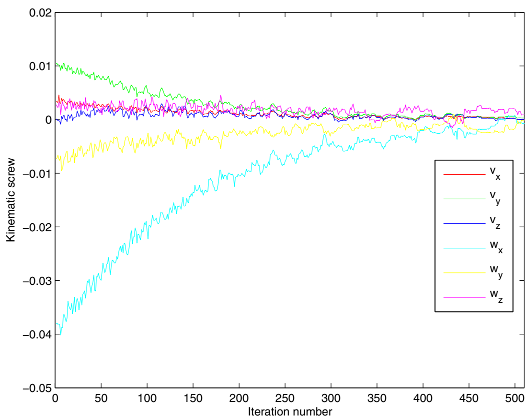

Fig. 5. Kinematic screw for reaching the pre-grasp position.

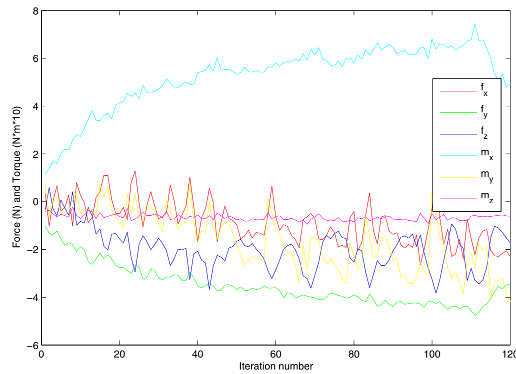

Fig. 6. Forces during grasping.

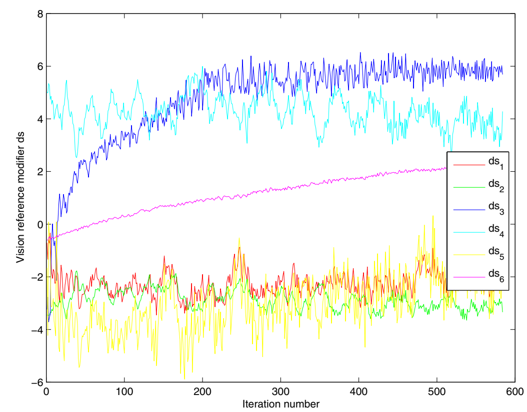

Fig. 7. Visual reference modifier based on iteraction forces.

2) Reaching the grasp position: During this step, the tool frame is moved, from the pre-grasp position towards the grasp frame, as shown in Figure 4c. A new vector of visual features is computed according to equation 1. Thus, the handle is reached from the reaching direction established by the transformation ${}^{\mathcal{P}}\mathbf{T}_{\mathcal{G}}$. At the end of this step, the grasp frame and the tool frame are the same (up to modelling errors), which means that the handle is situated between the robot fingertips.

3) Performing the grasp: At this step, the robot gripper is closed in order to grasp the handle, as shown in Figure 4d. During this step the first contacts appear. Thus, at the same time that the gripper is closed, the vision/force control law is active. The reference for the vision control law is to match the grasp and tool frames (i.e, keep the handle in the middle point between both fingertips). The reference for the force control law is to minimize external forces ($\mathbf{f}^d = 0$). If, due to modelling errors, the handle is not perfectly placed in the middle point between the fingertips, then one finger will make contact before the other. This will generate a force that the force control law will try to regulate to zero. During this step, the stiffness coefficient on Y direction of frame $\mathcal{E}$ is set to a small value in order to have compliance.

The real behavior is shown in Figure 6. Due to a premature contact on one of the fingers, it appears a force in Y direction and a torque in X axis (expressed in effector frame $\mathcal{E}$). These forces modify the visual reference (i.e. the grasp frame pose w.r.t tool frame) according to equation 3, and then the robot is visually-guided in order to reduce the force.

C. Interaction

Once the handle has been reached, the robot computes the direction of the force that must be applied at the contact point, depending on the motion that must be applied to the object [2]. Thus, the reference for the force control law $\mathbf{f}^d$ is set online, depending on the task (see Figure 4e). However, the reference for the vision control law is not modified, because the contact (and, thus, the relative position between object and gripper) must be kept constant during the task execution.

The new force reference will modify the original one, so that the robot will move in a direction suitable for the task guided by the vision law. The natural object mechanism will generate forces on the robot hand that the force control law will try to minimize, making the robot hand to adapt to the object motion. As the object pose is continuously observed, any misalignment between the hand and the handle will be detected and corrected. Thus, both force and vision will work simultaneously for a common goal: performing the task while the relative position between hand and handle is kept constant.

Experimental results on the interaction phase can be seen in Figure 7, where the evolution of the visual reference modifier $d\mathbf{s}$ is shown, which depends directly on the task forces according to equation 3. The visual reference is modified mainly in translation in Y and Z axis, and in rotation in X axis, due to the existence of important forces in these directions. Force in Z direction (of the end-effector frame $\mathcal{E}$) is regulated to a positive value according to the force reference $\mathbf{f}^d$, and it corresponds to the resistance of the particular object mechanism. The rest of forces appear on constrained directions and must be regulated to zero. The force in Y direction and torque in X direction are generated by the particular trajectory when opening the door. The force control law updates the vision reference so that the robot hand adapts to the natural trajectory. It is worth noting that the hand trajectory is never planified. Instead, the vision/force control law adapts the hand motion automatically to the particular object mechanism.

V. CONCLUSIONS AND FUTURE WORK

An integrated sensor-guided robotic manipulation system for common everyday tasks has been presented. The system combines a task-oriented grasp planning algorithm with advanced visual/force servoing capabilities. The task-oriented grasp planning module computes a grasp on the object taking into account the task to perform. An external position-based visual servoing approach is used in order to visually guide the hand of the robot towards the object to grasp. During this step, the robot's head is continously tracking the hand and the object. The relative pose between both is computed at each iteration independently of the camera position, which makes our approach amenable to be integrated into current humanoid robots without hand-eye calibration. Finally, the task is executed by means of a novel vision/force coupling approach which avoids control problems by making the integration in sensor space. Both vision and force feedback cooperate during task execution in order to keep the relative pose between gripper and object at the same time that the natural object mechanism is tracked. As future work, we would like to use the natural object features as input to the virtual visual servoing pose estimator. We would also like to work on head control algorithms in order to keep always a good view of the gripper and object during task execution. Task scheduling can also be improved for taking into account joint limits, obstacles, and other kind of task-relevant criteria by following, for example, the task sequencing approach recently published in [16]. Finally, we want to test the system on many different objects and mechanisms that future humanoid robots will have to deal with.

VI. ACKNOWLEDGEMENTS

This work is supported by the Intelligent Robotics Program, one of the 21st Century Frontier R&D Programs funded by the Ministry of Commerce, Industry and Energy of Korea. This work is also supported in part by the Science and Technology Program of Gyenggi Province as well as in part by the Sungkyunkwan University. And this work was also partly supported by Brain Korea 21 (BK21) project and by the Korean Ministry of Information and Communication. The authors want to thank the INRIA Lagadic team for the ViSP software [17].

REFERENCES

[1] Y. Mezouar, M. Prats, and P. Martinet. External hybrid vision/force control. In *Intl. Conference on Advanced Robotics (ICAR'07)*, Jeju, Korea, 2007.

[2] M. Prats, A.P. del Pobil, and P.J. Sanz. Task-oriented grasping using hand preshapes and task frames. In *Proc. of IEEE International Conference on Robotics and Automation*, pages 1794–1799, Rome, Italy, April 2007.

[3] R.P. Horaud, F. Dornaika, and B. Espiau. Visually guided object grasping. *IEEE Transactions on Robotics and Automation*, 14(4):525–532, August 1998. ISSN: 1042-296X.

[4] G. Taylor and L. Kleeman. Flexible self-calibrated visual servoing for a humanoid robot. In *Proc. of the Australian Conference on Robotics and Automation*, pages 79–84, Sydney, Australia, November 2001.

[5] E. Marchand and F. Chaumette. Virtual visual servoing: a framework for real-time augmented reality. In *EUROGRAPHICS 2002*, volume 21(3), pages 289–298, Saarebrücken, Germany, September 2002.

[6] K. Hosoda, K. Igarashi, and M. Asada. Hybrid visual servoing / force control in unknown environment. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1097–1103, Osaka, Japan, 1996.

[7] G. Morel, E. Malis, and S. Boudet. Impedance based combination of visual and force control. In *IEEE International Conference on Robotics and Automation (ICRA'98)*, volume 2, pages 1743–1748, Leuven, Belgium, May 1998.

[8] V. Perdereau and M. Drouin. A new scheme for hybrid force-position control. *Robotica*, 11:453–464, 1993. ISSN: 0263-5747.

[9] S. Lee, S. Lee, J. Lee, D. Moon, E. Kim, and J. Seo. Robust recognition and pose estimation of 3d objects based on evidence fusion in a sequence of images. In *Proc. of IEEE International Conference on Robotics and Automation*, Rome, Italy, April 2007.

[10] Ch. Borst, M. Fischer, and G. Hirzinger. Grasp Planning: How to Choose a Suitable Task Wrench Space. In *Proc. IEEE Intl. Conference on Robotics and Automation*, pages 319–325, New Orleans, LA, USA, April 2004.

[11] R. Haschke, J.J. Steil, I. Steuwer, and H. Ritter. Task-oriented quality measures for dextrous grasping. In *IEEE Conference on Computational Intelligence in Robotics and Automation*, Espoo, 2005.

[12] A.T. Miller, S. Knoop, H.I. Christensen, and P.K. Allen. Automatic grasp planning using shape primitives. In *Proc. IEEE Intl. Conference on Robotics and Automation*, pages 1824–1829, Taipei, Taiwan, September 2003.

[13] P. Martinet and J. Gallice. Position based visual servoing using a nonlinear approach. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1, pages 531–536, Kyongju, Korea, October 1999.

[14] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Trans. on Robotics and Automation*, 8(3):313–326, June 1992. ISSN: 1042-296X.

[15] E. Malis, F. Chaumette, and S. Boudet. 2 1/2 d visual servoing. *IEEE Trans. on Robotics and Automation*, 15(2):238–250, April 1999. ISSN: 1042-296X.

[16] N. Mansard and F. Chaumette. Task sequencing for high level sensor-based control. *IEEE Trans. on Robotics*, 23(1):60–72, February 2007. ISSN: 1042-296X.

[17] E. Marchand, F. Spindler, and F. Chaumette. Visp for visual servoing: a generic software platform with a wide class of robot control skills. *IEEE Robotics and Automation Magazine*, 12(4):40–52, December 2005.