Short survey

Dual arm manipulation—A survey

Christian Smith, Yiannis Karayiannidis, Lazaros Nalpantidis, Xavi Gratal, Peng Qi, Dimos V. Dimarogonas, Danica Kragic

Centre for Autonomous Systems, School of Computer Science and Communication, Royal Institute of Technology, 10044 Stockholm, Sweden

Article history:
Received 3 February 2012
Received in revised form 18 June 2012
Accepted 2 July 2012
Available online 15 July 2012

Keywords:
Dual arm robots
Domestic manipulation
Survey
Control
Modelling
Planning
Learning

1. Introduction

There is an increasing trend of robots being moved into environments originally designed for human use. In industry, anthropomorphic robots of human size are expected to replace human workers without major redesigns of the workplace. The ability to use human and robot workers interchangeably is thought to be the key to low-cost, flexible automation. As robot use in health care and domestic environments increases, so does the need for robots that are well adapted to these intrinsically human-centric environments. More advanced humanoid robots are envisioned to mimic human behaviors, act and manipulate objects in ways similar to humans.

This has, during the past few years, led to increased interest for the field of anthropomorphic or dual-arm manipulation. Robot manipulation in its basic forms is a well studied field that has seen remarkable developments in the last 50 years, but the added complexity of dual or multi-arm manipulation presents many challenges that may not be present in the single manipulator case. This higher complexity means that dual arm manipulation requires more advanced system integration, high level planning and reasoning, as well as viable control approaches. The challenges inherent in dual-arm manipulation, especially in unstructured environments, serve also as a motivator for development in basic research areas, and provide relevant application scenarios for various enabling technologies. Fig. 1 showcase the wide variation of dual arm systems.

The aim of this paper is to summarize recent developments in manipulator control, modeling, planning, and learning, with special attention given to work that specifically targets the dual arm case. There exist several reviews of each of the abovementioned domains in isolation, but the state-of-the-art and future challenges targeted specifically to dual arm cases are not easy to extract from these.

1.1. Background

Some of the very first robotic manipulators were dual-arm systems. Early examples include the manipulators constructed by Goertz in the 1940’s and 1950’s for handling of radioactive goods [1], that were used in pairs, with the operator controlling one with each hand. The late 1950’s also saw dual arm teleoperation setups for deep-sea exploration [2]. NASA’s Johnson Space Center started experimenting with anthropomorphic dual arm teleoperators in 1969 [3]. The history of dual arm manipulators and manipulation has been presented in detail in several earlier review papers [4–6].

This early work has been followed by an abundance of applications considering single arm manipulators. These have been a norm for a long time, especially in the 1980’s and 1990’s, when...
a lot of initial work on visual servoing was developed [11]. More recent developments in the areas of humanoid robots as well as the work on learning by imitation, has opened for use of dual arm setups. There are several independent factors that motivate the use of dual arm setups:

- **Similarity to operator**—The common motivation for using dual arms in teleoperation setups is that the operator is used to performing bimanual tasks, and having two arms at the slave site means that the operator’s bimanual skills can be transferred to the remote site [12–16].

- **Flexibility and stiffness**—By using two arms in a closed kinematic chain, the stiffness and strength of a parallel manipulator can be combined with the task flexibility and dexterity of a serial link manipulator [17].

- **Manipulability**—Dual arm manipulation is motivated by the ability to control both parts of for example a typical peg-in-hole task, with one arm positioning the peg and one arm the hole [18], or a screw assembly, where one arm controls the nut and the other the bolt [19]. The high degree of task space redundancy for dual arm systems has been used for optimal performance for domestic tasks such as dishwashing [20].

- **Cognitive motivation**—Human-like dual-arm setups have been used to explore how human-like physical interaction relates to cognition [21]. Likewise, in an HRI context, it has been argued that since humans have an intuitive understanding of bimanual manipulation, the actions of a dual arm robot are easier for an observing human to understand and relate to [22,18].

- **Human form factor**—With the advent of robots that are expected to work in environments originally intended for humans, it is claimed that robot manipulators will need to have a human-like form to perform well [23,24,8,25], even though some acknowledge that this may place significant constraints on performance – especially power and robustness – with contemporary hardware [26,22]. Recently, there have been several dual arm systems proposed for industrial use, with the motivation that dual arm systems occupy less space and have lower cost, as compared to two single arm units. Also the ability to replace human workers with robots without redesigning the workspace is used as a motivation [27–29,7,30].

This wide range of motivations for dual arm robots has led to the development of a large variety of hardware platforms, see Fig. 1 for a few examples. Even though the manipulator platforms may be very heterogenic, there are several common problems that have to be solved, as discussed in the following sections.

### 1.2. Definition

The term “dual arm manipulation” does not have a specific agreed-upon definition. Manipulation can be defined as physically interacting with an object, exerting forces on it in order to move or reshape it. However, “dual-arm” is not trivial to define. Two dexterous fingers mounted on the same hand may manipulate a small object (Fig. 1(e)) using the same principles as two separate manipulators that are moving a large object (Fig. 1(d)). In fact, many authors do not distinguish between multi-agent or multi-arm systems. A broad definition of cooperative manipulation is given instead, and covers the spectrum from different fingers on the same hand to teams of separate robots cooperating [31,32]. In terms of bimanual grasping, a detailed classification extending Cutkoskys grasp taxonomy for single-handed grasps to bimanual grasping has also been proposed [33].

One proposed general classification for dual arm manipulation makes distinctions between non-coordinated manipulation, where the two arms are performing two different tasks,
and coordinated manipulation, where the arms perform different parts of the same task. Coordinated manipulation is in turn divided into goal-coordinated and bimanual manipulation. For goal-coordinated manipulation, the arms are not physically interacting with one another, but both are solving the same task, with typing different keys on a keyboard given as an example. Bimanual manipulation is defined as physically interacting with the same object [34], see Table 1. Of these, the first definition delimits systems that may consist of two separate manipulators performing tasks independently of one another. As there is no explicit coordination, there is no intrinsic difference to single-arm systems, and the analysis does not need to differ from single-arm setups. The two later cases include significant amounts of spatial and temporal coordination aspects that set them apart from single-manipulator systems. The papers presented in the present review all belong to one of the two latter classes, and have been classified accordingly in Table 2. This table also classifies the papers according to their field of study, as a fast reference guide.

In the following section, we structure and review the state of the art in the area starting from modeling paradigms and finishing up with different application examples.

1.3. Dual arm systems

There exists a wide variety of robot platforms used for dual arm manipulation. While some research is carried out on systems built by simply placing two single-arm manipulators to share the same workspace [13,129,9], considerable effort has also been put into constructing dedicated dual arm platforms. Some of these put the effort on manipulation capability, and target industrial manufacturing applications, such as the Toyota Dual Arm Robot [132], the Yaskawa Motoman SDA10D [7], the ABB Frida [135] the Korea Institute of Machinery & Materials dual arm robot [28], the Kawada Hiro [29], the SHARP household robot [27], or the Pi4 Workerbot [30], each of which is shaped like a “torso”, and primarily intended for stationary deployment. Similarly, prototypical torso-type robots have also been constructed for different research purposes, like the Umass. Dexter [21].

A special class are torso-type robots intended to perform some of the tasks traditionally performed by human astronauts, like the NASA Robonaut [136] and CSA Dexter [137]. These can be mounted at the end of large manipulators, and have limited mobility.

Others put the emphasis on mobility, for operation in human environments. Examples include UMass. Ubot [138], Tohoku University Mr Helper [134] Willow Garage PR2 [8], DLR Rollin Justin [139,24], DFI Aila [140], KIT ArmarIII [22], TUM Rosie [141], Waseda Twendy-one [126], and GATech Domo [26] and Cody [142]. Yet others place the emphasis on completely mimicking human appearance and structure, such as the Honda Asimo [143] and the Kawada HRP Series [144]. A special class consists of robots especially intended for teleoperation work, where the appearance is not necessarily biomimetic, but the workspace is made to match that of the human operator as closely as possible [12–15,145,16].

1.4. Applications

While many dual-arm systems are pure research platforms mainly used for development and evaluation of basic technologies and principles, there are also several examples of practical applications. This section presents typical applications for dual arm manipulation (technical aspects of these systems are detailed in later sections).

As an example of domestic applications, folding laundry has been studied in several cases. These applications share the common aspect that they use vision to detect folds and corners, but differ in how the manipulation is performed. In [124], specialized “inchworm type” grippers on an anthropomorphic robot are used to physically trace edges. In [95] autonomous towel folding is demonstrated on the PR2. Stereo vision is used in conjunction with a vision-based grasp point detection algorithm that detects the corners of the towel, and a sequence of vision-based re-grasps and manipulations are performed towards the towel folding goal. In [122], the PR2 is used to grab one corner of a towel and lift it up and let gravity unfold it to detect further corners using vision. Early considerations about using a visually guided dual-arm robot to interact with clothes have been reported in [127], and in an industrial setting [129] focuses on fixed-frame vision-based techniques to detect and model the shape of clothing items to manipulate with industrial robots, and compare the performance of grabbing the highest point of a pile of towels with grabbing a rough approximation of one end of a towel. In [100] a dual-arm robotic setup for clothes manipulation using a pair of industrial robots (Kawasaki Js2 and Yamaha Rch-40) is presented. Images from a single fixed-position workspace camera are processed so as to detect whether a gripper has reached a corner.

Other domestic application types have also been described for dual arm robots. In [126], the TWENDY-ONE robot is presented. The robot is developed for providing attendant care and kitchen support for the elderly. In [123], a pair of dual arm anthropomorphic robots are used to make pancakes. The paper focuses on high level reasoning, planning and systems integration. In [27], a smaller dual arm robot is demonstrated to serve tea using preprogrammed motions and teaching by demonstration. Loading and unloading a dishwasher with the anthropomorphic robot ARMAR-III has been described in [22,125]. In [99] the work is extended with a visual servoing controller for dual arm grasping and manipulation tasks. The control framework allows it to track multiple targets. A prototype robot named RIBA with an anthropomorphic torso is proposed to perform health care related lifting tasks, such as transferring a human from a bed to a wheelchair [150].
In [96] a cooperative scenario between two 6-DoF dual-arm mobile robots is considered. The two independent robots grasp and move a box to a target position using visually estimated poses of the robots and the box. In [134], a control system of a mobile dual-arm robot for handling a single object in cooperation with a human is presented. The robot system performs cooperative manipulation tasks with a human, or transports the object autonomously.

In industrial settings, a typical application for dual arm systems is parts assembly. Preprogrammed gearbox assembly has been described in [132], while [28] describe programming by demonstration for similar tasks. Another area studied is material reshaping. In [131], a dual arm system is used to bend metal parts, and in [133], a dual arm system is used to fold cartons into predetermined shapes. Finally, [130] investigates the Learning-from-Observation paradigm for robot learning for industry-oriented assembly tasks for rigid polyhedral objects.

In [94], a dual-arm crew-assisting robotic system for space exploration missions is examined. The goal is the system to autonomously grasp objects using a vision and force-based control strategy. The system consists of a four-wheeled rover and two robotic arms, each with a camera mounted on the arm’s end-effector. An image-based visual servoing is used in conjunction with fiducial markers to move each of the robot’s arms.

Table 3 summarizes the different hardware platforms used in the papers cited in this work.

### 2. State of the art

The goal of this section is to structure and review the work in dual arm systems. We start by providing an insight and review the work on the modeling paradigms. This is followed by an overview of control theoretic approaches, where the aspect if vision based control is studied in more detail. Then, we review the work on planning and relation to grasping and manipulation of objects that consider dual arm setups. We also review different learning approaches and finish up by presenting a number of interesting applications.

#### 2.1. Modeling

The study of dual-arm robotic systems is an extensive research area and there is a vast literature concerning how to model the system to achieve grasping and manipulation of an object or an object-independent target. Part of this section is based on textbooks [151,152,32,43] and review papers on robotic grasping and manipulation [153,154]. This section presents literature regarding the characteristics of closed chain systems, such as the grasp matrix, the internal forces, the restraint properties, the force/load distribution, the inherent redundancy as well as the implications of manipulation of deformable sheets and the use of mobile manipulators in dual arm manipulation.

The modeling of a robotic system which consists of multiple cooperative manipulators employed to manipulate an object, heavily depends on the type of the interaction between the end-effectors and the manipulated object. There are two main categories of cooperative manipulator tasks:

1. Cooperative robot manipulators holding an object with fixed grasp points (Fig. 2(a)): The object is assumed to be rigidly attached to both manipulators, and hence no relative motion...
the grasping point to a frame \{o\} attached at a fixed point of the object (typically the center of mass) can be defined as follows:

\[
G = \begin{pmatrix}
I_\text{p}_{oc1} & 0_3 & -I_3 & 0_3 & I_3 \\
-S(p_{oc1}) & I_3 & -S(p_{oc2}) & I_3 & I_3
\end{pmatrix}
\]

where \(p_{oc}\) is the relative position of frames \{c_i\} and \{o\}, \(O_3, I_3\) are \(3 \times 3\) null and identity matrices respectively and \(S(\cdot)\) is a skew-symmetric matrix being used in order to produce cross-products. The virtual-stick, which is defined for each manipulator to denote the relative position between the end-effector and a fixed point of the object, can be regarded as the core of the kineto-static formulation [32]. On the other hand, the cooperative task-space formulation directly defines the task space variables as absolute and relative motion of the dual arm manipulation system, can be applied even for tasks which do not require physical interaction of both arms with the object [152]. In case of contact points (contact areas) the contact modeling is crucial to derive the grasp matrix; in this case, \(G\) given by (1) can be characterized complete grasp matrix, while the inclusion of selection matrices for the components of forces/torques transmitted through the contact yields the grasp matrix.

In particular, the contact modeling determines the components of the end-effector wrenches at the contact point which are going to reach the object and thus the components of the velocity which are transmitted to the object. There are three main type of contact models [151,153]:

- **Frictionless point contact** is used to model cases wherein very small contact areas arise allowing the slippage between the contacted parts.
- **Frictional point contact or hard finger model** approximates the cases of small area contact able to transmit tangential frictional forces. When very small contact areas arise the friction moment around the contact normal is negligible.
- **Soft contact or soft finger model** is used when the fingertip or end-effector material is soft enough and thus capable of producing significant contact areas. An arbitrarily large contact area enables fingertip to exert frictional torques around the contact normal.

Soft contact may intuitively seem more practical, but is not commonly utilized; details on soft contact modeling for dexterous manipulation can be found in [48].

The type of contacts described above do not affect modeling of dual arm manipulation with fixed grasps but it is important for cooperative manipulation in cases of the end-effector cannot achieve a fixed grasp. The grasping formation as well as the type of contact (in the case of non-fixed grasp) discussed above are used in order to derive the constraints that governs the manipulation task. The Lagrange multipliers associated with differentiable algebraic
constraints have been used in order to model the forces in dual arm manipulation arms with no relative motion [36]. Nonholonomic constraints are imposed in case of rolling motion between the manipulators and the contacted object [44]. In [37] the dynamic modeling of a dual-arm manipulator using constraints has been presented along a comparison between experimental and simulation results. In case of manipulators with soft end-effectors, like for example soft fingers, the normal interaction force is the resultant of the spring forces, and consequently it is related to the material deformation and to the deformation velocity in the case of deformable viscoelastic materials [63]. Hence, for compliant contacts the interaction forces are functions of the system state while for rigid contacts they depend on both the system states and the system inputs [38]. Modeling of frictional contacts for the fingertips manipulating an object typically obeys the Coulomb law and thus the magnitude of the tangential frictional forces are proportionally related to the pressing forces [151]. The replacement of the sign function for Coulomb law by smooth functions is proposed in [63]. The friction coefficient determines the cone of friction and is subsequently directly involved in the slippage detection condition. The part of the exerted force which does not contribute to the motion of the object is the so-called internal or squeeze forces. Recently, port-Hamiltonian modeling for object manipulation by soft fingers has been proposed in order to describe uniformly the transition from the non-contact case to the visco-elastic contact case [62]. The virtual linkage [39] and virtual truss models [49] have been used in order to characterize the range of the internal forces in multiple cooperative manipulators, for the cases of a tightly grasped object and of an object manipulated by fingertips.

In the case of dual arm manipulation using fixed grasp points, the restraint properties such as force or form closures are not of interest. In the case of dual arm manipulation using contact points or contact areas, the force and form closure properties have been extensively studied in the literature. A survey of the state-of-art before 2000 can be found in [153]. The difference between the form and force closure properties lies in the contact model that is assumed. The force closure properties rely on the ability of the contact to exert frictional forces which are dependent on the exerted normal force. The form closure property is more restrictive since it relies on a frictionless contact model. In addition, passive and active closure properties have been introduced in [50] to characterize the ability of the grasping mechanism to transfer motions to the grasped object.

Han et al. formulate the grasp analysis problems as a set of convex optimization problems involving linear matrix inequalities, and show the simplicity and efficiency of this LMI formulation [51]. The choice of optimal grasp synthesis and its evaluation, as well as the force distribution problem for object grasping [52,159,153], are of fundamental importance and have been extensively explored in recent publications [53–56].

The problem of force distribution for an object manipulated by fingers is analogous with the load distribution problem for cooperative constrained manipulators. The load distribution problem is related to finding ways to share the load to each arm according to its actuation ability [152]. New perspectives in the grasp synthesis arise by considering the concept of “postural synergies” or “hand configuration subspace” implying an underactuated system of cooperative manipulators [57,58].

Recently, there has also been significant interest in understanding how to manipulate deformable objects such as rope, napkins and cloth [35]. Wakamatsu et al. [60] analyze stable grasping of deformable objects based on the concept of a bounded force closure, and apply the proposed method to linear deformable objects. Modeling of deformable objects is an active research area and different approaches have been considered. A useful survey can be found in [160]. In robotics, low-complexity representations for the deformable objects should be the objective. One simple and recently well used representation for the deformation manipulation is to consider the deformable object as a mechanism, with creases as revolute joints and facets as mechanism links. This has been successfully exploited in automatic paper folding [161] and industrial carton handling [162].

In dual arm manipulation, redundancies may arise even in cases where two single arm manipulators with no inherent redundancies are used. The redundancies are a direct result of the cooperative task definition. The controlled task space variables are 6 generalized positions while internal forces also can be controlled. Redundancies are also useful for increasing the robotic hand or the cooperative dual-arm manipulation system dexterity. The kinetostatic performance indices such as velocity and force manipulability ellipsoids have been formally expanded to evaluate performance quality for cooperative arm manipulation [40] and mechanical hands with cooperative fingers [59]. The effect of softness at the contact area on the manipulability and the grasp stability of a system which consists of multiple manipulators grasping an object have been recently studied in [61].

The modeling of dual arm manipulators systems becomes more complex when the robot arms are mounted on mobile or free-flying platforms. Specifically, the modeling of mobile manipulators holding a deformable object is considered within the framework of Kane’s approach in [41]. The dynamic equation of motion for dual free-flying arms are described in [42].

2.2. Control

In this section we review the main results on both dynamic and kinematic control approaches considering manipulation of an object using cooperative manipulators. Thus, we provide an overview of important works of the literature, in which the underlying solution is designed by closing the loop at the operational space and/or the interaction force level; in this case the operational space is mainly defined based on the object position, while in some cases task coordination positions are used. We here exclude task hierarchical levels used to achieve the same objective by combining higher level grasp and motion planning (producing the desired joint trajectories) with lower level joint space control that drives the robot to the specific configuration. Section 2.1 as well as the survey [153] describe several higher level grasping and motion planning techniques specific for dual arm setups, while standard control techniques for joint position regulation and tracking for single arms can be used without any modification. Hence, this section mainly treats papers that consider force/motion control and closed chains (either flexible or rigid) formed by the manipulators and the object. In robot control literature there are two main categories of problems: regulation and trajectory tracking. Such categorization can be adopted for dual-arm manipulation control to distinguish between the cases of a constant and a time dependent desired object pose and internal forces.

The trajectory tracking problem has mainly been treated using input–output linearization methods which are in general model-based, where the estimation step requires the knowledge of the robotic system structure. Nonlinear feedback techniques have been used to linearize the dynamic equations of the closed chain motion [38]. The constraints imply that one part of the output, which is the interaction forces, depends on the input; thus input integrators have been utilized. An input–output linearization systematic approach has also been used for 3D rolling contacts [44] and for point contact with friction and compliance [64]. The application of input–output linearization for cooperative manipulation through compliant contacts enables the
control of the normal force without employing input integrators, but increases the required design and implementation effort by giving rise to the singularity of the decoupling matrix.

Hybrid force/position control and impedance control have been extensively used in dual-arm manipulation tasks. For a single arm manipulation setup, hybrid force/position control is based on the decomposition between the motion and exerted force control loops, while the impedance control method simultaneously controls the motion along every direction in order to achieve a desired dynamic interaction between the manipulator and the environment. The basic structures of hybrid force/position and impedance control for dual-manipulator systems manipulating an object which may be in contact with the environment as well as a comparison of their performance can be found in [45]. An impedance controller may lack precision in controlling the position of an object as compared to hybrid force/position control but can be applied without any switching procedure between contact and non-contact cases.

Hybrid force/position control for dual-arm manipulation tasks is mainly based on the decomposition of the object motion and the internal forces space using kineto-static filtering and thus enables the direct control of internal forces which is crucial in many dual arm manipulation applications. In [65], the authors propose independent nonlinear feedback with hybrid force/position control of each robot of the dual-robot setup; the controller is driven by reference trajectories which are obtained by a centralized planner aiming to achieve coordination for a pushing and pulling operation. Force/position hybrid control has been proposed in [66] in order to control the internal forces at a desirable level and to enable load distribution of the control effort simultaneously with the main task of object position tracking; cooperative tasks with respect to rigid objects such as screwing a nut onto a bolt have been considered. A feedback linearizing input and the grasp matrix have been exploited in [67] in order to decompose the object motion control directions and the internal force directions. This work in addition proposes a hybrid control for rolling and rolling/sliding contacts. Exploiting the orthogonality between the spaces of active and passive force closure, a decomposing control input which enables the simultaneous control of internal forces and object motion is proposed in [68]. Adaptive passivity-based hybrid force/position control has been designed in the joint space by incorporating the law of Slotine-Li for dynamic parametric uncertainties [36] and the need of velocity measurement in such a type of controllers has been studied in [69]. Decomposition of the force and position control loop is also used in a model-based control scheme that requires derivatives of force measurements in order to control a cooperativeness error which is related to the velocities of the arms lying on the normal to the object directions. The corresponding error can then be regarded as a measure of pushing/pulling forces, [70]. Hybrid force/position has been also combined with vibration suspension control in order to control a dual-flexible arm system manipulating a rigid object, [71]. Moreover, hybrid control has been utilized for underactuated cooperative manipulators—in which case recalculation of the dynamic load-carrying capacity is proposed in order to verify the ability of the manipulator to execute the cooperative task [72].

Schneider and Cannon [78] have initially defined the desired impedance for the object dynamics in contact with the environment enlightening the cooperative perspective of impedance control. The generic structure of an impedance controller requires the acceleration of the object in order to be implemented, but estimates as well as finger force sensing can be used instead, [45]. Furthermore, the impedance controller can be used for internal force control as described in [79,80]. Recent research has proposed also the combination of a decentralized impedance controller for each manipulator with centralized impedance control for the manipulated object [9]. The last results in the use of impedance control in dual arm manipulation can achieve a desired impedance for the motion of the object with respect to its environment as well as desired impedance for the motion of its manipulator with respect to the object (indirect internal force control). Both impedance and hybrid force/position tracking control are based on the robot and object dynamic model.

The ideas of sliding mode control, robust adaptive control as well as neuro-adaptive control have also been used for cooperative manipulator tasks in order to cope with dynamic model uncertainties by assuming a closed chain model. Intelligent control using neural networks and fuzzy systems have been proposed in [73,87] respectively in order to cope with dynamic model uncertainties. Neuro-adaptive control with visual feedback has been also proposed in order to deal with Jacobian kinematic uncertainties [74]. Such types of techniques are straightforward generalizations from the one to multiple manipulator case, since dynamic model uncertainties affect the dynamic response of the controlled system in a similar way in both cases.

In case of regulation, the entire dynamic model of the robotic system, however, is not required and hence simpler control structures have been proposed. In the last decade, effort has been put into designing regulators for the grasping and manipulation of an object by a pair of robotic fingers; this effort can be transferred to the context of dual arm manipulation. In [10] two rigid fingers have been used to control the object position and orientation simultaneously, with rigid object grasping. The stability proof for this is based on passivity and proposes the concepts of stability and transferability on a manifold in order to cope with redundancies. The aforementioned concepts do not require the definition of additional tasks and augmented projection method for task prioritization as in [81]. Several extensions on this work have been done towards (i) control of whole arm and dual arm with fingers for achieving grasping and manipulation ([82,47] respectively), (ii) 3D object motion [46], and (iii) multiple fingers manipulating an object [10]. The implementations of the aforementioned controllers use estimates of the orientation and position of the objects in order to avoid the use of vision, but position-based visual servoing can also be used. Simple controllers accompanied by passivity-based stability proofs have also been proposed for the case of soft fingers manipulating rigid materials [83,84]. The softness of the fingertip in general enhances dexterity but also requires measurements of the fingertip deformation and some additional control terms, specifically in the case of deformable fingertips [84]. Passivity-based regulation for three robotic fingers manipulating an object in the 3D space with 3D deformation have been considered for grasping and manipulation [75]. Object-level impedance behavior without inertia shaping has been applied to multi-finger hands for dynamic dexterous manipulation [158] and extended to the dual-arm manipulation problem [25].

The aforementioned work considers an object-centered workspace. On the other hand, there is work which considers the cooperative task–space formulation for fixed-grasp typical dual arm manipulation problems and proposes very simple control laws based on PD with gravity compensation structure. The two basic approaches are the following: (i) finding the desired joint values that correspond to the cooperative task using inverse kinematic algorithms and substituting them within a joint space PD control scheme [76], and (ii) finding the desired position of each manipulator based on the desired absolute and relative position of the task–space formulation and substitute them within the robot’s operational space PD controllers [77]. The regulators can be also combined with internal force feedback and can be regarded as hybrid force/position controllers since they exploit the kineto-static filtering of the control actions. Recently, kinematic control for performing dual-arm manipulation using dual quaternions has been:
proposed [88]. The dual quaternions have been used to define the "cooperative dual task-space" since they simultaneously describe the position and the orientation of a rigid body.

The perspective of control of dual- or multiple-arm manipulation systems via the coordinated motion of multiple autonomous agents mainly considers robot arms mounted on mobile platforms and is an ongoing research topic. In this case additional issues such as obstacle avoidance and exchange of information between the manipulators need to be taken into account. Kinematic control for nonholonomic manipulators grasping a deformable object has been applied in order to move between two states in a known static environment with obstacles without over-stretching the object [89]. Navigation functions with an obstacle avoidance term have been exploited to achieve the control objective. The tracking of fixed-base, free-floating, and free-flying co-operating manipulator systems have been treated by using inverse dynamics control algorithms with motive force [42]. Schemes with one agent acting as a leader and one or more agents acting as followers have been considered for the coordinated motion of multiple mobile manipulators holding a single object [85]. The leader knows the desired trajectory of the object while the followers estimate it and manipulate the object by coordinating with the leader. Synchronization of the manipulators has been addressed in [86] where an adaptive synchronized controller is proposed to drive the synchronization error to zero and hence maintain certain kinematic relationships between the cooperative manipulators. The work in [86] has been extended in [163] to accommodate a control loop for the internal forces under both flexible and rigid constraints. However, the proposed direct force cancelation technique cannot be applied in general.

2.3. Vision and visual servoing

Visual feedback has been widely used for providing robots with the ability to perceive and interact with their environment [23]. Vision based techniques were originally developed for single-armed robots and have been later adapted for use in dual-arm setups. As a result, explicitly dual arm-oriented vision systems are not commonly found in the relevant literature. In practice, there are two different ways for visual information to be exploited. The first one is by establishing a vision based closed-loop position correcting procedure, a technique which is known as visual servoing. The second one is by observing the environment once, calculating a desired movement and then applying it without any further corrections.

The task of visual servoing is to establish a closed control loop for a robot to manipulate objects within its environment using visual feedback [11]. The concept and the techniques of visual servoing have been described by Hutchinson et al. in 1996 [164] and more recently by Chaumette and Hutchinson in [165,166]. Visual servoing techniques for dual-arm robots are usually extensions of the techniques used in the simpler case of single robot arms. As such, they can be also classified into the same categories, i.e. image-based, position-based, and hybrid methods.

Image-based visual servoing uses the difference between the observed and desired position of selected features on the 2D image plane. As such, no pose estimation is required and the calculations that provide the robot control signals can be performed very rapidly. As an example, an image-based visual servoing technique is presented in the work of [74] that uses the image Jacobian matrix to control multi-fingered robot hands. Furthermore, the use of planar markers as the guide for image-based visual servoing has been exploited in [94].

On the other end of the spectrum, position-based visual servoing uses the difference between the observed and desired pose of the tracked object in the 3D space with respect to the camera. A 3D model of the object is reconstructed based on some extracted features and its pose in the Cartesian space is calculated either by exploiting information of a pre-captured model (for monocular vision systems) or by taking into consideration depth information as well (for visual sensor arrangements providing depth). Within this context, the anthropomorphic robot ARMAR-III has been used to perform grasping of objects considering the possible configurations of both the left or the right arm [98]. The same platform has been used in [99] to demonstrate a position-based visual servoing controller for dual arm grasping and manipulation tasks. Furthermore, [92] also uses a position-based visual servoing scheme for the control of two robot manipulators, and [19] uses the end-effector mounted camera on each arm to align the parts held by each hand to perform an assembly task.

Visual sensors can also provide general information about the context or the elements of a scene. Such information could be exploited in ways different than for controlling the arms, i.e. for understanding a scene, or localization purposes. As a result, active vision has to do with observing the environment and making decisions about the robot's next actions. However, this kind of decision-making is not a closed loop control scheme.

Visual systems can be characterized according to where their point of view is placed. Most of the proposed systems employ one (or possibly a combination) of the following camera configurations:

- **End-effector mounted**—This configuration, often also called eye-in-hand configuration, has the camera mounted on the robot's end-effector. This configuration provides a predefined geometric relationship between the position and orientation of the camera with respect to the arm. A typical example of such a configuration in a dual-armed robot are the visual sensors placed on the end-effectors of each of the Willow Garage PR2 robot's arms [8].
- **Fixed in the workspace**—This configuration provides a constant, stable place for the vision system to observe the scene.
- **Active head**—Active heads provide the vision system with a limited flexibility of translating and rotating with respect to the arms' reference system.

These configurations are graphically depicted in Fig. 3.

The first two configurations, depicted in Fig. 3(a) and (b), are more commonly found in industrial and generally non-domestic robotic setups. The case of an end-effector mounted vision system has been studied in [90]. The work of [94] examines a robotic system for space exploration missions that has two arms, each with a camera mounted on the end-effector and uses an image-based visual servoing strategy. Recently, the popular PR2 robot has also adopted, apart from an active head, an eye-in-hand camera configuration with limited reported results yet. The configuration of two cameras fixed in the workspace is used in the work of [92], which is further discussed in [93], for vision-based control of two industrial arms. Finally, the system presented in [97] also uses a fixed vision system to control a robot having two industrial arms.

On the other hand, the active head configuration, depicted in Fig. 3(c), is a common choice for anthropomorphic robots found in domestic environments, due to its similarity to the human vision apparatus. An example of such a setup is the ARMAR-III robot [125], whose vision system is comprised of an active head and a double binocular camera setup for wide and narrow angle vision, respectively. Furthermore, the work of [126] presents the TWINDEY-ONE elderly assisting robot that employs a 3-DOF active head for its stereo vision system. Finally, the active head of the PR2 robot and its vision system have been used for vision-based towel folding in [95].

Another, characterizing factor of vision systems is the kind of visual sensors they employ. The most commonly used systems use
uses a Randomized Path Planner to generate a trajectory for the transfersubpaths, which together form a manipulation path, and 3D environment. It again breaks the problem into transit and problem of several robotic manipulators moving an object in a version of the problem. In [107], the same authors consider the paper presents successful experimental results for a very simple and the object remains static) and transfer (where the arms move independently, order to move it. In order to avoid collisions between arm, object and obstacles, the arms may have to change their grasp of the object. The method proposed extends the randomized potential field planning techniques introduced in [116] with the introduction of transit (where the arms move independently and the object remains static) and transfer (where the arms and the object form a closed kinematic chain) subpaths. The paper presents successful experimental results for a very simple version of the problem. In [107], the same authors consider the problem of several robotic manipulators moving an object in a 3D environment. It again breaks the problem into transit and transfer subpaths, which together form a manipulation path, and uses a Randomized Path Planner to generate a trajectory for the object from its initial to its goal configuration. The Randomized Path Planner algorithm [116] is modified to ensure that for each configuration along the path, not only is the configuration collision-free, but also the object can be grasped. The set of transfer paths is thus generated. Then, RPP is used again to generate the transit paths between the different transfer paths. The method is tested with 3-6-DOF arms.

Manipulability ellipsoids are defined as the mapping of the unit hypersphere at the origin of the joint velocity space into an ellipsoid in the Cartesian velocity space by a Jacobian transformation. These manipulability ellipsoids can be used to describe, at certain critical task points, the requirements for Cartesian motions and static forces. This allows the definition of a task-oriented manipulability measure (TOMM), which expresses the similarity between the desired and actual manipulability ellipsoids at the selected task points. In [117], this concept is extended to define a task-oriented dual-arm manipulability measure (TODAMM), which is the similarity between the desired manipulability ellipsoid and the intersection of two individual manipulability ellipsoids.

An approach to constrained motion planning for robotic systems with many degrees of freedom is introduced in [108]. It first establishes which are the conditions under which the manipulation constraints are holonomic. Then, it simplifies the problem, which in general includes non-holonomic constraints, so that all the constraints are holonomic. This is done by replacing an equality-constrained problem with a converging series of inequality-constrained problems, penalizing motions which do not satisfy the constraints. The result of each step of the series is used as the input for the next step, and each of the problems can be solved using a standard path planner.

In [109], a very specific problem arising in space applications of dual-arm systems is described. It studies the case of a free-floating dual-arm manipulator, where one of the arms is required to perform a certain task, while the other performs the necessary compensating motions to keep the base inertially fixed. The main problem faced is to avoid the singularities. It uses algorithm based position kinematics equations together with an iterative search procedure. The study shows the potential of the method, while finding configurations where the methods fail.

LaValle in [110] proposes a formulation of the multiple-robot planning problem, and describes several strategies for decoupled planning:

- Prioritized planning assigns priorities to each of the robots in the system. Then, the path is calculated for the first robot without considering the rest. Then, the method is repeated for each of the other robots, considering only the higher-priority robots for collision detection.
- Fixed-path coordination considers the case where each robot must follow a fixed path, and the goal is to find the timings that allow the robots to do so in a way that the robots don't collide with each other. Examples of this strategy are described in [111,112]
- Fixed-roadmap coordination extends fixed-path coordination to enable each robot to move over a roadmap or topological graph

Segsin et al. [113,114] introduce control configuration points (CCPs) for using redundancy for obstacle avoidance when using multiple planar robot manipulators. CCPs are selected along the kinematic structure of a manipulator, usually in the midpoint of links. These are used by Liu and Dai [118] to develop a method for planning the trajectory of multiple robotic fingers for carton-folding. They create a path connection based on CCPs and geometry guiding points (GGPs). The method generates the minimum number of fingers required for a certain manipulation of the carton, and integrates the carton motion into the trajectory of the robotic fingers.

Based on the RRT algorithm first introduced by LaValle in [115], Vahrenkamp et al. introduce in [119] a method that combines into a single planner the three tasks needed for grasping an object: finding a feasible grasp, solving the inverse kinematics and finding a trajectory from some initial position to the grasping pose. A single probabilistic planner based on RRTs is used to search for a grasp that is both feasible and reachable, without the need of using precalculated grasping positions, which limit the possible grasps to a predefined set. The paper presents experimental results for bimanual grasping with the ARMAR-III robot. ARMAR-III has also been employed to demonstrate the application of probabilistically complete RRT-based motion planning algorithms for solving the problems of inverse kinematics, dual-arm manipulation motion planning and re-grasping tasks [167].

Also based on RRTs, the method in [120] introduces the concept of task maps, which represent the manifold of feasible grasps for an object. The task maps are learnt using the RRT algorithm. Then, choosing a grasp and choosing the path from the starting position of the robot to the position where the object is grasped is treated as a single problem, which leads to choosing a grasp which is easy to reach from the current position of the robot. The problem is solved using a gradient-based optimization method.

As an alternative to roadmap and RRT-based planning, [121] introduces a new planning method based on a new approach for representing the state. This is done based on models of the robot and its environment, which aim to take advantage of the structure found in robotic environments, and uses multiple variables to describe the current state. These variables may represent body parts, objects or constraints, but also provide abstractions for the current state. They assume that the state of the system is described by these random variables.

2.5. Learning

To achieve the goal of having robots that can work in unrestricted environments, it is necessary to equip them with the ability to adapt to the surroundings. Without the ability to learn, human designers must manually program all aspects of the behavior of the robot, which is burdensome in unrestricted environments, where it is impossible to foresee every possible situation that the robot might encounter. Learning by demonstration has been used for some time to help with this learning process, by enabling robots to learn new actions by looking at a human performing them. Most of these methods have been traditionally aimed at a single manipulator. While it is possible, in some cases, to trivially extend these methods to two arms, just by repeating the process for each arm, there are some peculiarities of dual-arm manipulation which require some modifications to the framework [101,168].

In [103], a framework for dual-arm programming by demonstration (PbD) is described. Bi-manual actions are classified based on the spatial relationship between the trajectory of hands performing coordinated actions. This leads to a threefold classification into coordinated symmetric, coordinated asymmetric and uncoordinated actions. This classification is made from the recorded demonstration, which is divided into one and two-hand fragments. Two-hand fragments are those where both hands are in the “grasped” state. Coordinated symmetric action is then detected by looking at closed kinematic chains. Several heuristics are then used to distinguish between asymmetric and uncoordinated action. For the execution of the actions, a synchronization framework based on Petri nets is implemented.

The system in [104] uses Hidden Markov Models and Gaussian Mixture Regression to deal with the variability across several demonstrations and extract redundancies. Though it is not specifically dual-arm based, it has been successfully used to learn a bimanual dancing motion containing crossings, on an iCub robot.

In [105], a framework is presented that aims at teaching a bimanual coordination task to a robot. This method combines Dynamical Systems movement control with a Programming by Demonstration approach. The Programming by Demonstration method is a low-level one, which encapsulates the task at the trajectory level. Their main contribution is the use of a dynamical system approach which is used to encode coordinated motion patterns, which allows the use of a PbD approach to learn the movements of both arms simultaneously, and focusing on the relation between them.

A new framework for the learning of coordinated movements is presented in [102]. The system builds on the concept of learning control as learning a policy which maps the state to a control vector. It uses nonlinear dynamic systems as policy primitives. These primitives, called programmable pattern generators (PPG), implement a globally stable attractor with specifics determined by the values of a set of parameters. Different types of PPGs can represent different types of movement relationships. They implement their system on a 30-DOF Sarcos Humanoid Robot, using both arms to generate a regular rhythm on a couple of drums.

3. Where are we heading?

A future direction in robotic systems in general, and in dual arm manipulation in particular, will be the integration of elements from systems theory with tools from cognitive methodologies. These will involve the consideration of vision and learning capabilities in the actual feedback design. This seems necessary in advanced collaborative control tasks where the manipulators have incomplete knowledge of the environment, and might have been assigned their tasks independently. In particular, the lack of global knowledge about the environment and the final objective of the task in hand, along with the limitations in coordination due to sensing and communication constraints, suggest the integration of cognitive features in order to modify the individual agent controllers in an on-line data-driven manner. Future research will thus build, among others, towards the integration of the different research features addressed in this paper in order to address complex coordinated tasks.

Examining sensing and perception, this survey has shown a multitude of studies detailing different approaches, and several application papers where a specific sensor configuration is more or less given by technical capabilities or limitations of the chosen platform. In future work, it would be of interest to have a deeper discussion on the choice of sensors. Not only do we need
to ask what sensors we should use, what modalities they should have, or where they should be placed, but we also need a discussion on what motivations should be driving these choices. Is it meaningful to mimic the (proven successful) configuration of biological systems, such as humans? Should more advanced sensors be used to cover deficiencies in planning, modeling, and control, or should the latter be developed to cover deficiencies in sensors?

We could also turn the question around, and examine how developments in one subsfield influences another. For instance, how would developments in sensing and perception affect modeling and control? What developments in sensing and perception would be necessary to support desired developments in planning?

How should we model and implement high level understanding of environments, objects and tasks. How can we support high-level reasoning about known entities? How do we extrapolate from existing knowledge in order to fill in gaps? How do we transfer domain-specific knowledge from humans to robots?

There are several possible routes towards answering these questions. We will need methods to analyze human behavior based on information from vision, haptics, or other sensors, along with reasoning systems to determine what aspects of a task are relevant, and need to be replicated by the robot. We will need strategies to map the human task performance to a dissimilar manipulator, identifying the goals of a task rather than the exact procedure. We will need efficient planning systems to achieve these goals.

With dual arm robots working in domestic and other human-centric environments, we also anticipate a substantial increase in the number of collaborative robots. All of these issues need to be addressed before robots can successfully replace humans in manipulation tasks in unstructured environments.

Acknowledgments

This work has been supported by the Swedish Research Council (VR), the Swedish Foundation for Strategic Research (SSF), and the European Union FP7 project RoboHow.Cog (FP7-ICT-288533).

References


[120] M. Gienger, M. Toussaint, C. Goerick, Task maps in humanoid robot


[118] S. Lee, Dual redundant arm configuration optimization with task-oriented

[117] J. Barraquand, J.-C. Latombe, Robot motion planning: a distributed representa-


[115] U. Sezgin, L. D. Seneviratne, S. W. E. Earles, Redundancy utilization for obstacle


Christian Smith received his Master of Science in Engineering Physics from the Royal Institute of Technology (KTH) in Stockholm in 2005, and his Ph.D. in Computer Science from the same institution in 2009. He has been a postdoctoral researcher at the Intelligent Robotics and Communication Laboratories at ATR in Kyoto, Japan, and is currently a post-doctoral researcher with the Centre for Autonomous Systems at KTH. Research interests include domestic manipulation, humanoid robot interaction, and human–robot interfaces.

Lazaros Nalpantidis is currently a post-doctoral researcher at the Centre for Autonomous Systems (CAS), Computer Vision & Active Perception Lab. (CVAP) of the Royal Institute of Technology (KTH), Sweden. He holds a Ph.D. (2010) from the Department of Production and Management Engineering, Democritus University of Thrace, Greece in the field of Robot Vision. He received a B.Sc. degree (2003) in Physics and an M.Sc. degree (2005) in Electronic Engineering both from the Aristotle University of Thessaloniki, Greece. His current research interests include active and cognitive vision systems for robots in scene understanding, depth perception, obstacle avoidance and SLAM applications.

Xavi Gratal started his Telecommunications Engineering studies at Universitat Politècnica de Catalunya in Barcelona, and completed his Master Thesis on "Visual Servoing" at the Royal Institute of Technology in Stockholm, as part of a double degree program. He is currently pursuing his Ph.D. at the Royal Institute of Technology. His research interests include Visual Servoing and Motion Planning.

Dimos V. Dimarogonas received the Diploma in Electrical and Computer Engineering in 2001 and the Ph.D. in Mechanical Engineering in 2006, both from the National Technical University of Athens (NTUA), Greece. He has held postdoctoral positions at the ACCESS Linnaeus Center, Royal Institute of Technology (KTH), Stockholm, Sweden and the Laboratory for Information and Decision Systems (LIDS) at the Massachusetts Institute of Technology (MIT), Boston, MA, USA. He is currently an Assistant Professor at the School of Electrical Engineering, Royal Institute of Technology (KTH), Stockholm, Sweden. His research interests include Multi-Agent Systems, Hybrid Systems and Control, Robot Navigation and Networked Control.

Yiannis Karayannidis was born in 1980. He is currently a postdoctoral researcher in Computer Vision and Active Perception Lab. at Centre for Autonomous Systems at the Royal Institute of Technology (KTH) in Stockholm, Sweden. He received a diploma in Electrical and Computer Engineering in 2004 from the Aristotle University of Thessaloniki, Greece and a Ph.D. from the same University in 2009, working on the control of kinematically uncertain robotic systems under a grant from the State Scholarships Foundation of Greece. His research interests are in the area of robot control.

Danica Kragic is a Professor at the School of Computer Science and Communication at KTH in Stockholm. She received an M.Sc. in Mechanical Engineering from the Technical University of Rijeka, Croatia in 1995 and a Ph.D. in Computer Science from KTH in 2001. She is the Director of the Centre for Autonomous Systems at KTH and the coordinator for the EC integrated project GRASP and TOMSY. From 2011 she has been a member of the Royal Swedish Academy of Sciences and Swedish Young Academy. Danica received the 2007 IEEE Robotics and Automation Society Early Academic Career Award. Her research is in the area of computer vision, object grasping and manipulation and human–robot interaction.