

PAPER

Stereovision and augmented reality for closed-loop control of grasping in hand prostheses

To cite this article: Marko Markovic *et al* 2014 *J. Neural Eng.* **11** 046001

View the [article online](#) for updates and enhancements.

You may also like

- [Weighted total least squares–Bayes filter-based estimation of relative pose for a space non-cooperative unknown target without a priori knowledge](#)
Chengguang Zhu, Zhongpai Gao, Jiankang Zhao *et al.*
- [A two-step method for spatial circle orientation with a structured light vision sensor and error analysis](#)
Bin Wu, Ting Xue and Shenghua Ye
- [A novel optical method for measuring 3D full-field strain deformation in geotechnical tri-axial testing](#)
Pengpeng Wang, Yong Sang, Xiaoxia Guo *et al.*

Breath Biopsy Conference



Join the conference to explore the **latest challenges** and advances in **breath research**, you could even **present your latest work!**



5th & 6th November
Online



Main talks



Early career sessions



Posters

Register now for free!

Stereovision and augmented reality for closed-loop control of grasping in hand prostheses

Marko Markovic¹, Strahinja Dosen¹, Christian Cipriani², Dejan Popovic³ and Dario Farina^{1,4}

¹ Department of NeuroRehabilitation Engineering, Bernstein Focus Neurotechnology Göttingen, Bernstein Center for Computational Neuroscience, University Medical Center Göttingen, Georg-August University, D-37075 Göttingen, Germany

² The BioRobotics Institute, Scuola Superiore Sant'Anna, V.le R Piaggio 34, I-56025 Pontedera, Italy

³ Faculty of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11020 Belgrade, Serbia

E-mail: marko.markovic@bccn.uni-goettingen.de, strahinja.dosen@bccn.uni-goettingen.de, ch.cipriani@sssup.it, dbp@etf.rs and dario.farina@bccn.uni-goettingen.de

Received 19 October 2013, revised 7 February 2014

Accepted for publication 4 March 2014

Published 3 June 2014

Abstract

Objective. Technologically advanced assistive devices are nowadays available to restore grasping, but effective and effortless control integrating both feed-forward (commands) and feedback (sensory information) is still missing. The goal of this work was to develop a user friendly interface for the semi-automatic and closed-loop control of grasping and to test its feasibility. **Approach.** We developed a controller based on stereovision to automatically select grasp type and size and augmented reality (AR) to provide artificial proprioceptive feedback. The system was experimentally tested in healthy subjects using a dexterous hand prosthesis to grasp a set of daily objects. The subjects wore AR glasses with an integrated stereo-camera pair, and triggered the system via a simple myoelectric interface. **Main results.** The results demonstrated that the subjects got easily acquainted with the semi-autonomous control. The stereovision grasp decoder successfully estimated the grasp type and size in realistic, cluttered environments. When allowed (forced) to correct the automatic system decisions, the subjects successfully utilized the AR feedback and achieved close to ideal system performance. **Significance.** The new method implements a high level, low effort control of complex functions in addition to the low level closed-loop control. The latter is achieved by providing rich visual feedback, which is integrated into the real life environment. The proposed system is an effective interface applicable with small alterations for many advanced prosthetic and orthotic/therapeutic rehabilitation devices.

Keywords: artificial proprioception, closed-loop control, control of grasping, stereovision, augmented reality

 Online supplementary data available from stacks.iop.org/JNE/11/046001/mmedia

(Some figures may appear in colour only in the online journal)

⁴ Author to whom any correspondence should be addressed.

1. Introduction

Reaching and grasping is a complex movement during which the human neuromusculoskeletal system accomplishes a number of very challenging tasks [1]. The process starts with the visual assessment of the target object to determine its location and orientation (required to correctly reach it) as well as its properties (shape and size required for prehension). This information is used by the brain to plan the motor task [2]. During the actual movement execution, the hand is transported to the target object and at the same time oriented and preshaped appropriately for the grasp based on the visual perception. The tactile and force feedback and proprioceptive feedback are used for online correction, especially during the contact phase. Importantly, these complex sensory–motor transformations are implemented by the neuromusculoskeletal system with remarkable ease and almost effortlessly for the subject, i.e., the subject is mostly concerned with the overall goal of the task (e.g., use of the object), while the actual implementation proceeds largely at a subconscious level (automatically). This natural control is made possible by a system of sophisticated sensory organs and human effectors in communication with the trained brain through a complex network of efferent and afferent channels, implementing the feed-forward and feedback sensorimotor schemes [3].

The grasping function, both motor and sensory aspects, can be lost or impaired as a result of a neurological injury, such as stroke, spinal cord injury, brachial plexus injury or limb amputation. Depending on the context, different assistive devices have been developed to restore grasping (e.g., functional electrical stimulation (FES) [4], exoskeletons [5] or prostheses [6]). Importantly, many of these devices are very advanced technological solutions striving to replicate the flexibility of the healthy human hand. Multichannel electrical stimulators with programmable timing and intensity can activate several muscles of the forearm in a coordinated fashion generating different grasps [4]. In recent years, several hand exoskeletons including separate kinematic chains for each finger have been developed and tested for force amplification and/or rehabilitation [7]. Similarly, modern prosthetic hands such as i-Limb from Touch Bionics [8] or Bebionic Hand from RSLSteeper [9] implement independent control of individual fingers. These robotic systems have enough mechanical flexibility to realize different grasp types, reflecting closely the repertoire used in daily life by a healthy human hand (e.g., palmar, lateral, pinch grip). However, user friendly control of grasping is still missing, and thereby the full capabilities of these advanced systems are largely unexploited in practice. Typically, only one grasp type is used [7] or the active grasp has to be selected manually through an unnatural (unintuitive) interface (e.g., pressing a button in FES [4], cocontracting antagonistic muscles in prosthetics [10]).

The control of grasping should include both a higher level interface for planning (grasp type, hand orientation and size selection) and a low level control of grasp execution, ideally providing also sensory feedback to the user (closed-loop control). Effective and effortless implementation of these high and low level functions is a long standing challenge in

rehabilitation engineering, with the most extensive research done in the field of prosthetics. In a typical approach, the control is realized by implementing a unidirectional communication pathway between the user and artificial system. The user generates all the necessary command signals (inputs), whereas the artificial controller captures the inputs, decodes them and operates the device accordingly. Therefore, the user is in control of the operation, and there is no automatic function that would decrease the need for large number of control inputs. Different signal sources have been tested to generate the control signals. In principle, high level motor commands for grasping can be detected directly at their origin (brain) using electroencephalography [11] or electrocorticography [12], but this approach is still far from being practical enough for use in daily life and requires complex instrumentation. In the previous studies, grasp type and size selection (high level control) were implemented by using foot movements detected by an insole equipped with foot switches [13], tongue control via a mouth piece integrating an inductive interface [14] and eye movements detected using electrooculography [15]. By far the most prevalent approach relied on the detection of the user's muscle activity [10, 16]. Pattern recognition was applied to classify hand postures and even individual finger movements from the recorded multichannel surface electromyography (EMG) with high classification accuracies [17–19]. Also, implantable interfaces such as intra-fascicular neural electrodes [20] and implantable myoelectric sensors [21] have been tested for the same purpose. In contrast to a large body of work in recent years addressing the descending motor pathways, there are significantly fewer studies devoted to the restoration of feedback to the user. However, this was a popular research topic in the past [22, 23], and direct mechanical (e.g., vibrotactile or haptic [24, 25]) or electrocutaneous stimulation [26] of the tactile sense were investigated as the means for providing the feedback. In some recent works, vibration motors [27], custom-made devices [28], electrical and hybrid stimulation [29] and also invasive approaches [30] have been tested.

Although success has been demonstrated, the establishment of the communication between the biological and artificial remained the challenge to be resolved. The main reason is that conventional man–machine interfaces used in practice have a limited bandwidth in both descending (feed-forward) and ascending (feedback) directions. Therefore, they cannot fully support the available functionality of the modern-day assistive devices. More recently, there are developments presented in the robotics literature inspired by human motor control demonstrating that a low-dimensional control input can be used to implement complex movements [31] and also facilitate learning [32]. A myoelectric controller presented in [33] implemented hand synergies to operate a dexterous hand prosthesis (three grasps) using only two myoelectric channels. These important advances should be used in the future to reduce the demands upon the user efforts regarding how to use the assistive system, and instead allow him/her to concentrate on the task to be accomplished.

We suggest here an approach for decreasing the burden on the user regarding how to use the assistive device.

The idea is to enrich the artificial controller with extra sources of information, as also recently proposed in [34], in addition to the classically used human-generated command signals (e.g., myoelectric control), so that the controller is capable of autonomous decision making [35]. Consequently, certain system functions could be controlled fully automatically while the role of the user would be to supervise, correct and fine tune the system operation. The idea to automatize the assistive devices or facilitate the control using additional information sources has been addressed in the past. An early example is a prosthesis implementing automatic grasp selection based on the place of initial contact with the target object [36–38]. A commercially available prosthetic device, Sensor Hand Speed from Otto Bock [39], and some other recently presented systems [40, 41], integrate controllers that automatically compensate for object slipping. In [42] and [43], hierarchical control strategies in which myoelectric signals were used to trigger predefined high level functions (e.g., activating specific grasps) were developed and tested. Finally, an interesting sensor fusion approach, although for a different purpose, was presented in [44] where an eye tracker estimated the gaze direction, and this information was then combined with the myoelectric control to greatly improve reaching using a robotic arm.

We present a novel system for the semi-automatic and closed-loop control of grasping for assistive devices for the restoration of grasping function. As explained above, visual information is truly instrumental for the planning and execution of grasping. Therefore, we developed an artificial controller that employs state-of-the-art computer stereovision to emulate the human visual sense and automatically configure the hand into a preshape pattern that is appropriate for the target object. The fundamental components of the system are augmented reality (AR) glasses with embedded stereo cameras. The AR interface was employed to provide the user with augmented and intuitive visual feedback about the status of the hand for online correction of the grasp (i.e., closed-loop control). In particular, the AR interface is used to (virtually) project the hand into the field of view, thus informing the user about hand preshape even when it was out of view (artificial proprioception).

Automated grasp planning and execution using vision but also other interfaces such as Kinect have been addressed previously in the literature, especially in the context of autonomous robotics [45]. However, the system presented here is unique since it is designed for a specific context of human–machine interaction. Namely, it integrates high-level functions of an automatic artificial controller with the volitional (biological) control of the user into a robust control loop operating online. The main aspect of this development is that the operational responsibility, and thus the cognitive load, are both shared between the system and the user.

The system as a whole, as well as some of the concepts and components, is being applied for the first time in this context (e.g., AR glasses, AR feedback, user-controller bidirectional communication, fusion of EMG and artificial vision). Therefore, the goal of this study was to present the control system and its components and then to evaluate

the overall feasibility of the proposed approach. We tested the control system with healthy subjects who used a dexterous robotic hand in a simple manipulation task. Specifically, the goal was to address and evaluate: (1) the performance of the fully automatic as well as the semi-automatic control; (2) the user ability to successfully operate the system when sharing the control with the artificial controller; (3) the feasibility of the AR feedback by assessing if the user is able to properly perceive and successfully utilize the information provided by the AR interface. The results demonstrate that our approach is effective for the control of assistive devices, and the method becomes practical as the technologies that facilitate the method are developing rapidly and are convenient for daily use (e.g., Google glasses [46]).

2. Material and methods

2.1. Overall control structure and system components

Figure 1 summarizes the operation of the system and shows how the tasks and responsibilities are shared among different functional blocks. The user wears AR glasses with embedded stereo cameras. He/she triggers the operation of the semi-autonomous controller via a simple myoelectric interface. The semi-autonomous controller selects grasp type and aperture size appropriate for the target object by processing the visual information retrieved by the stereo cameras. Accordingly, the controller sends a preshaping command to the hand. The selected grasp type and hand aperture size are visually fed back to the user through the AR glasses (AR objects). The user can employ the AR feedback to adjust the automatic decisions through the myoelectric interface. Eventually the controller implements hand closing to grasp and hand opening to release the object, in response to user myoelectric commands. The presented control is best described as semi-automatic since the user triggers the automatic operation but can also override/correct all of the decisions of the autonomous controller.

The system comprises the following components: (1) a standard laptop computer (8GB RAM, i5@2.73 GHz) equipped with a USB data acquisition card (DAQ 6210, National Instruments, US), (2) AR glasses (iWear920AR, Vuzix, UK), (3) a two-channel analogue myoelectric (EMG) amplifier (anEMG2, OTBioelettronica, IT) and (4) a robotic hand prosthesis prototype (IH2 Azzurra, Prensilia, IT).

The robotic hand was a left-handed, commercial version of the SmartHand [6]. It consists of four fingers and a thumb actuated by five electrical motors. Allowed motions are flexion/extension of the thumb, index, middle and ring–little finger as a couple and the rotation of the thumb opposition space. The hand includes encoders on each motor and an electronic controller that implements position control, by receiving commands sent over a serial bus from the laptop. The hand is mounted on a custom splint made from thermoplastic material and strapped firmly using Velcro straps and elastic bands to the subjects' forearm, so that the robotic hand is positioned directly beneath and parallel to the subjects' hand. Four different grasp types were implemented: palmar, lateral, bidigit pinch and tridigit pinch grasp.

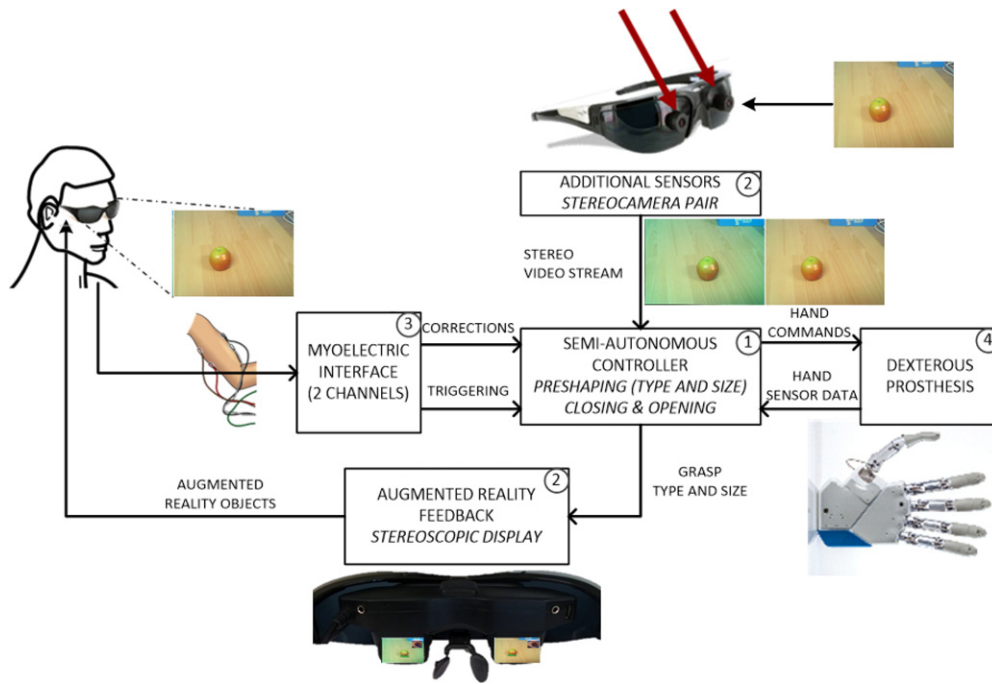


Figure 1. System architecture. The system comprises the following components: (1) semi-autonomous controller implemented on a standard laptop equipped with a data acquisition card, (2) AR glasses with an integrated pair of stereo cameras, (3) two-channel myoelectric interface and (4) a multi-grasp robotic hand.

Each of the two channels of the EMG amplifier includes a second-order band pass Butterworth filter (10 and 500 Hz). The amplifier gain is adjusted individually for each subject to obtain good signal-to-noise ratios (visual inspection). Two pairs of standard Ag/AgCl EMG electrodes (Neuroline 720, AMBU, USA) are placed over the left-hand flexor and extensor muscles, with a reference electrode on the wrist. The amplifier outputs are acquired by a data acquisition card (1 kHz sampling rate) connected to the laptop.

AR glasses are worn by the subjects. The glasses include stereo cameras embedded into the glass frames (resolution: 640×480 pixels, refresh rate: 30 Hz). The cameras are therefore directed towards the scene the user is looking at. The video stream acquired from each camera is sent to the laptop via a USB port. The image frames are processed and then projected via a VGA port at the two panels embedded in the glass frames. The panels implement a stereoscopic display and therefore allow the user to see a single 3D image of the scene in front of him/her (i.e., as if he/she is looking without the glasses). The scene contains both real and virtual objects (AR feedback).

2.2. Control flow and algorithm implementation

The control flow is implemented as a finite state machine in which transitions between the states are triggered by a simple two-channel myocontrol interface. A high-level representation of the state machine is shown in figure 2(a) and illustrated using snapshots in figure 2(b) and a detailed algorithm is given in table A1 in the appendix.

The user could start operating the system by directing his/her sight and thus the glasses to the target object

(figure 2(b), panel 1). After recognizing the target object, the system would acknowledge to the user the recognition of the target in the form of an overlay covering the object (object selection AR feedback) (figure 2(b), panel 2) and the user could then issue the command for opening the hand. This action would trigger the semi-automatic control and the system would analyze the scene (computer vision) to determine the geometrical model and properties (size and shape) of the target object. Based on this analysis, the system would estimate the grasp type and aperture size appropriate for grasping the object and send a preshape command to the hand (figure 2(b), panel 3). The component for stereovision analysis and subsequent grasp type and size selection based on the artificial vision is named stereovision grasp decoder. The AR feedback about the automatically selected grasp type and aperture size would then be presented to the user (figure 2(b), panel 4). The grasp-type AR feedback is presented as a visual icon, i.e., an image in the upper right corner of the subject's field of view. The aperture-size AR feedback is presented in the form of a virtual box placed just next to the target object, with one side parallel to the support plane and the other side aligned with the front side of the object. The size of this AR box is proportional to the current hand aperture size as read from the hand sensors. The user could therefore assess whether the hand was appropriately sized to grasp the target object by simply comparing the size of the virtual box to the size of the actual target object. The user could exploit this feedback to adjust the aperture (e.g., aperture too small/large) or to restart the decision process from the beginning (e.g., wrong grasp type) by issuing specific EMG commands. If the user judged the grasp type and aperture as correct, then he/she could bring

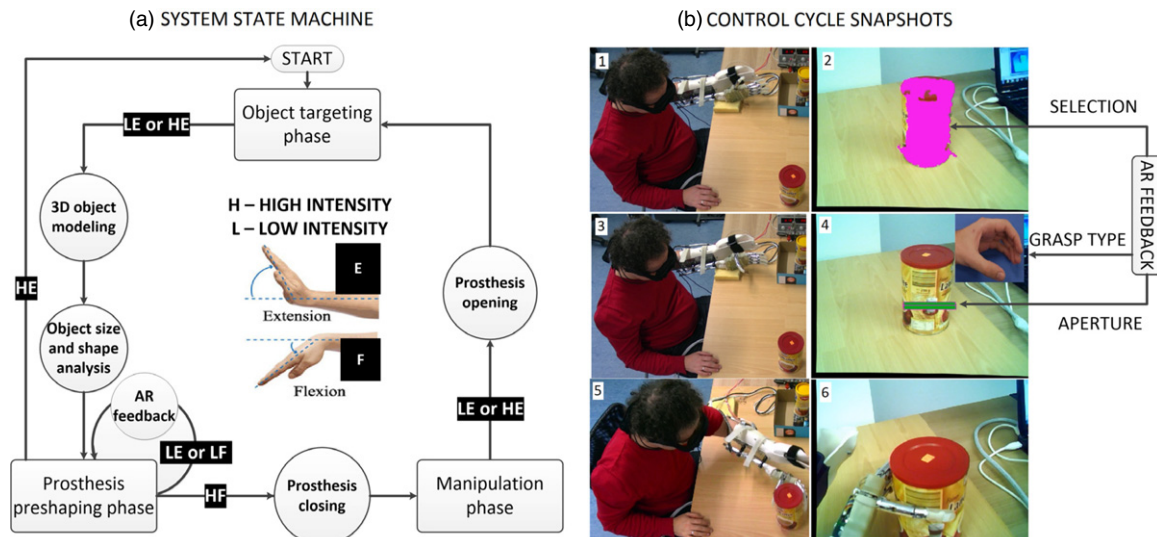


Figure 2. The control workflow: (a) system state machine and (b) control cycle in snapshots (left panel shows the setup and right panel depicts what the user actually saw through the glasses). The control loop operated as a state machine with the state transitions triggered by a simple two-channel myoelectric interface (high/low \times flexion/extension). From top to bottom, the snapshots depict: (1) object targeting phase with AR feedback about object selection, (2) hand preshaping phase with AR feedback on the selected grasp type and aperture size and (3) hand closing phase.

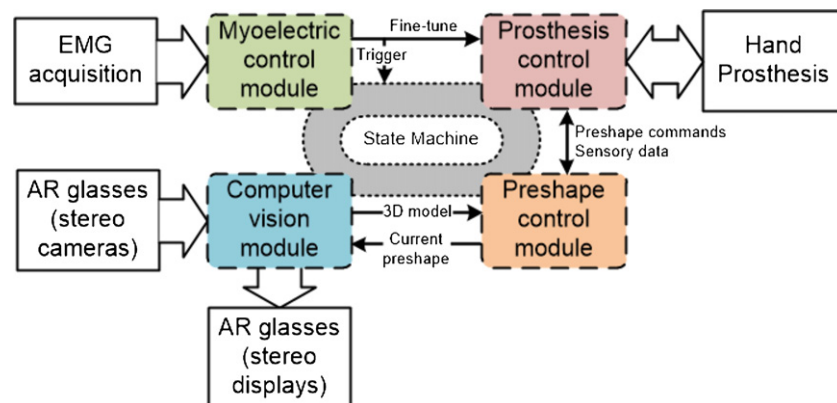


Figure 3. The interconnection of the system control software components and data flow between them. The online control software was implemented in Matlab and organized in the form of several modules with well-defined inputs and outputs.

the hand into the object vicinity and issue the EMG command to close it (figure 2(b), panels 5 and 6), manipulate the object as wished and finally reopen the hand in order to release the object (and restart). The described control system was running online and the average time needed to respond to a preshape command was less than 1 s (using the equipment described in the previous paragraph). This time includes command detection, stereovision processing and sending of the preshape commands to the hand. The online control software running on the laptop was implemented in Matlab 2012 and organized as a set of individual modules shown in figure 3 and described in the following. For more detailed information, see the supplementary data document attached to this paper (available from stacks.iop.org/JNE/11/046001/mmedia).

Myoelectric control module. This module implements the online acquisition and processing of the EMG signals.

The inputs for the processing are the EMG signals picked-up from the flexor and extensor muscles of the hand. The output is a 4-bit code indicating four possible user-generated commands (flexion/extension \times high/low). Four commands are generated by using two-level thresholding of the RMS calculated over a 200 ms window (with 50% overlap) of raw EMG, and ‘the first signal wins’ strategy was used to resolve the eventual co-activation. This 4-bit code is fed into the state machine (figure 3) to trigger the state transitions and into the prosthesis control module to adjust the grasp properties (type and aperture size) (figure 3). Figure 4 demonstrates how myoelectric processing was used to command the system throughout the whole operation sequence (figure 2).

Computer vision module. This module is responsible for the bidirectional communication with the AR glasses, i.e., acquisition of the images from the stereo cameras

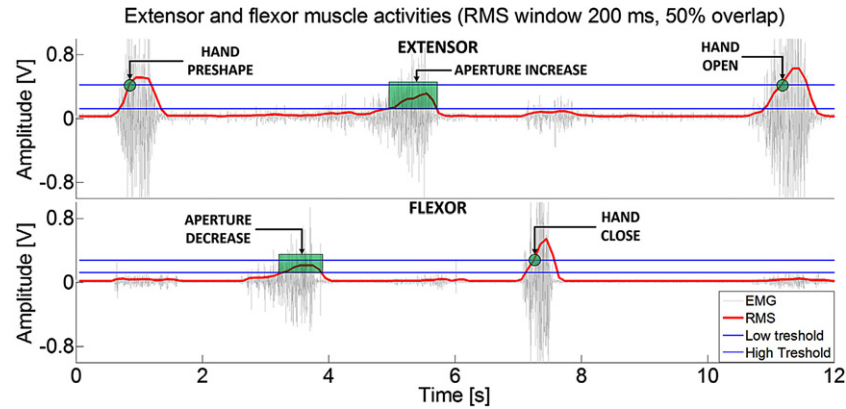


Figure 4. An example of myoelectric control during a complete trial. Top and bottom plots show the extensor and flexor activity, respectively. The thresholds for low and high activation are depicted using horizontal lines. Note that the thresholds are different for the two channels. Initially, the subject generated a high-level extensor contraction to open the hand. In response, the system initiated computer vision processing and automatically preshaped the hand. The subject then fine-tuned the hand aperture by generating low level contractions of the flexor (decrease aperture) and extensor (increase aperture). Finally, high flexor activation triggered the automatic hand closing, and thus the object was grasped. The trial ended with the automatic hand opening in response to a high level extensor contraction.

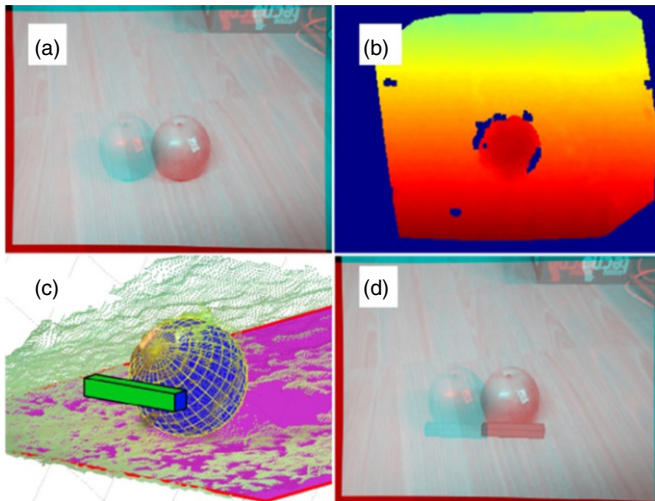


Figure 5. Illustration of the computer vision processing pipeline (computer vision module): (a) stereo pair showing the target object (an apple), (b) depth image (warmer colors denote the pixels closer to the viewer), (c) geometrical model (sphere) fitted through the cloud of points representing the target object and AR feedback object (green box) inserted into the 3D image, and (d) AR object projected into the pair of stereo images that are sent to the stereoscopic display within the AR glasses.

and control of the stereoscopic displays, and also for the processing of the image data. The inputs for processing are the two images from the stereo camera pair, and the current preshape of the hand (grasp type and aperture size) supplied by the preshape control module. The first input is used to estimate the properties of the target object (control) and the second to construct the AR feedback object (sensory feedback). The module implements a state-of-the-art computer vision processing pipeline comprising the following steps (figure 5): (1) depth estimation (figures 5(a) and (b)), (2) object segmentation, (3) 3D point cloud generation, (4) fitting of a

geometrical model (box, cylinder, sphere and line) through the point cloud (figure 5(c)), and (5) construction of the virtual objects (AR feedback) and their projection into the real scene (figures 5(c) and (d)). The outputs from the module are the target object properties (shape and size) and stereo images with the embedded AR feedback. The size of the object was given in centimeters. The object properties are input for the preshape control module, whereas the images are sent to the stereoscopic panels of the AR glasses to be presented to the user. Depth estimation is implemented using an efficient large-scale stereo matching (ELAS) algorithm, as described in [47], which is applied on the rectified grayscale image pair. The cloud point is computed from the disparities using triangulation in combination with the extrinsic and intrinsic properties of the camera system, which are identified during calibration [48]. The RANSAC algorithm [49] with application specific modifications is used for point cloud processing and object modeling.

Preshape control module. The inputs for this module are the object properties estimated by the computer vision module and the current state of the hand preshape provided by the prosthesis control module, whereas the outputs are the automatically selected grasp type and size. The grasp type was selected using cognitive-like processing represented by a set of IF-THEN rules similar to the ones described in [50] and [35], while the grasp size was a continuous variable calculated as the object size estimate plus heuristically adopted extra margin of 0.5 cm. In essence, the rule base implemented a mapping from the object properties into an appropriate grasp type (i.e., IF OBJECT SHAPE is X and OBJECT SIZE larger/smaller than Y THEN use GRASP TYPE Z). Palmar grasp was used for wide cylindrical and spherical objects, lateral for thin box objects, tridigit and bidigit for smaller objects etc. The computer vision module together with the preshape control module constitutes the stereovision grasp decoder.

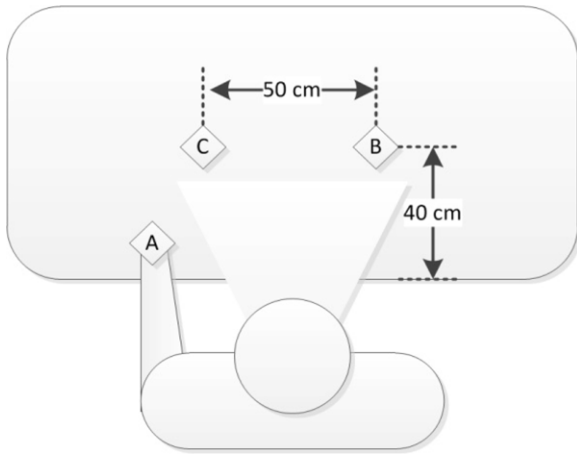


Figure 6. The experimental setup. The subjects were seated comfortably in front of a desk with three positions marked: (A) initial position to rest the prosthetic hand mounted on the subject's forearm by using a custom made splint, (B) initial position for the target objects and (C) final position to which the object had to be transported by the subject. The triangular area radiating from the subject towards the desk depicts the subject's field of view. Note that the hand was outside the subject's view when placed in the initial position.

Prosthesis control module. This module is responsible for the low level communication between the system and the hand prosthesis (i.e., sending commands and receiving sensor data over the serial bus with a sample rate of 20 Hz). The inputs are desired grasp type and aperture size (automatic control) and the user commands for the correction of the hand aperture (manual control) and the output is the current preshape state determined by reading the prosthesis sensors. The prosthesis actuators are position driven. The position commands that should be sent to the hand in order to implement a grasp of certain type and size are specified using lookup tables. The latter are prepared beforehand by measuring the response of the hand to position commands covering the available range of motion.

2.3. Experimental protocol

The control system was tested by able-bodied subjects performing a simple grasping task. Thirteen healthy subjects (29 ± 4 yr) volunteered in the experiments after signing the informed consent form approved by the local ethics committee. The subjects were comfortably seated on an adjustable chair in front of a table where a workspace was organized (see figure 6). Three positions were marked on the workspace: the initial (rest) position for the hand, the initial position for the target object and a position to which the objects had to be transported after being grasped. The workspace was organized and the system was mounted so that the robot hand was outside the user's field of view while he/she was looking at the target object.

At the beginning of the session, the subject was given instructions about the system operation and experimental protocol and then he/she was allowed to briefly familiarize with the system (less than 10 min). Twenty objects of different sizes and shapes (see table 1) were used as the targets and

Table 1. Tested objects, their sizes and corresponding grasp types.

Object	Size (cm)	Grasp type
Apple	6.8	Palmar
Coffee box	10×16.1	Palmar
Tea box	$6.5 \times 7.5 \times 16$	Palmar
Mug	8.2×9.7	Palmar
Deodorant	4.4×17.2	Palmar
Tennis ball	6.5	Palmar
Thick pen	1.8×13.8	Lateral
Thin pen	1×14.3	Lateral
Fork	1×19.4	Lateral
Thin DVD box	$0.8 \times 12.5 \times 14$	Lateral
Thick DVD box	$2.5 \times 14 \times 17.3$	Lateral
Small plastic brick	$0.7 \times 2 \times 4$	Bidigit pinch
Large bottle cap	3.5×2	Bidigit pinch
Small bottle cap	2.6×1	Bidigit pinch
Chestnut	3	Bidigit pinch
USB stick	$1.3 \times 2.6 \times 6.1$	Tridigit pinch
Big plastic brick	$3 \times 3 \times 4$	Tridigit pinch
Espresso cup	4.9×5.3	Tridigit pinch
Lip crème	1.9×6.7	Tridigit pinch
Crème box	1.8×5	Tridigit pinch

were presented to the subjects in a randomized order. Subjects were asked to reach, grasp, transport and release each of the target objects by operating the artificial hand using the semi-autonomous control loop explained in the previous paragraphs. The subjects were instructed to place the hand back in the initial position after the completion of each trial. Subjects operated the hand within four different experimental conditions with 5–10 min rest between the conditions.

- (1) *Automatic control with AR feedback (AUTO-AR).* In this condition, the subjects could not correct the decisions made by the system/stereovision grasp decoder (refer to table A1 in the appendix). Two series of 20 trials were performed. This condition was used to assess the baseline performance when the control was fully automatic. The AR feedback did not play a useful role in this condition but was displayed to the subjects so they could familiarize with it.
- (2) *Semi-automatic control with AR feedback (SEMI-AR).* In this condition, the system operated according to the full control sequence (without restrictions). Therefore, in addition to triggering the automatic control the user was able to correct the system decisions (grasp type and size) by relying on the AR feedback (refer to table A1 in the appendix). The subjects could decide if and how much to correct. They were instructed neither to look at the hand nor to move it from the initial position until the corrections were completed. This meant that all the adjustments had to be accomplished by relying exclusively on the AR feedback. The test included one series of 20 trials and the goal was to compare the performance of semi-automatic control with respect to the fully automatic control (AUTO-AR).
- (3) *Semi-automatic control with AR feedback and random errors (SEMI-AR-RE).* The control was the same as in the

Table 2. Outcome measures.

Class	Name	Description
GC1 ^a	Size estimation error	Absolute difference between the actual size (D) and estimated object size d : $GC1 = d - D $
GC2	Hand aperture success rate	The per cent of trials in which the selected hand aperture s was: $s > (D - 0.5 \text{ cm})$ (otherwise, the hand was not open enough to grasp the object)
GC3	Grasp type success rate	The per cent of trials in which the selected grasp type matched the correct grasp type for the given object as defined in table 1.
GC4	Preshape success rate	The per cent of trials in which both the grasp type and aperture size were correct (as defined above). In this case, the hand was perfectly preshaped to grasp the target object. Thus, GC4 was considered as the primary measure of performance.
TA1 ^b	Task accomplishment success rate	The per cent of trials in which the task was successfully completed (the user grasped the object, lifted it off the table, brought it to the final position and released it).
TA2	Grasp control task failure rate	The per cent of trials in which the user failed the task due to an incorrectly selected grasp type or size.
TTA	Time needed for the task accomplishment	The time interval from the moment when the first hand preshaping command was issued till the moment when the first closing command occurred.

^a GC stands for grasp control.

^b TA stands for task accomplishment.

SEMI-AR condition but the computer vision module was programmed to introduce random errors ($\pm 2\text{--}5 \text{ cm}$) in the estimated hand aperture size. This error was added only after the grasp type was determined. As a result the estimated grasp type was (mostly) correct but the hand aperture was surely wrong forcing the subject to readjust the hand based on the visual information provided by the AR feedback as in the SEMI-AR condition (i.e., without looking at the hand). Again, it was up to the subject to decide when and how much to correct. This scenario included two series of 20 trials and the goal was to test the general perception and usefulness of the AR feedback by forcing the subjects to rely on it while finely adjusting the aperture size in practically every trial.

- (4) *Semi-automatic control with random errors and without the AR feedback (SEMI-VIS-RE)*. The system operated identically as in the SEMI-AR-RE condition but with the AR feedback turned off. The subjects now had to bring the hand close to the object and correct the aperture by directly comparing the hand to the object. This condition was tested in one series of trials and was used as the control condition for the SEMI-AR-RE.

2.4. Data analysis

The correct automatic control of the prosthesis and the task accomplishment were correlated but not in an absolute sense. Indeed, the task could be accomplished although a wrong grasp and/or size were selected (e.g., lateral grasp to pick up a bottle or hand fully opened to grasp a small object). To take this into account, we adopted two sets of outcome measures (see

table 2), the first to evaluate the performance of the automatic or semi-automatic control of grasping (GC1–4) and the second to assess the actual accomplishment of the task (TA1–2). Additionally, in order to further quantify the stereovision grasp decoder performance and compare different operating modes, we have also measured the time from the moment when the first hand preshaping command was issued till the moment when the first closing command occurred, which is in the further text referred to as the time needed for the task accomplishment (TTA, see figure 4). Note that $100\% - (TA1 + TA2)$ represents a per cent of the cases in which both grasp type and size were correct but the user still failed to complete the task. Put differently, this outcome measure includes the trials in which the automatic or semi-automatic grasp control was successfully accomplished, but the subjects failed the task afterwards, due to some other reasons. This could be, for example, an unintentional myoelectric triggering while approaching the object or a slippage of the object from the hand due to the prosthesis construction. We grouped these cases together into a common category of myoelectric/prosthesis control failure. As explained above, although this measure does reflect the overall performance of the conglomerate system, it is not related to its core features, which are automatic and semi-automatic grasp control. In the case of the semi-automatic control (SEMI-AR, SEMI-AR-RE, SEMI-VIS-RE), the outcomes GC2–4 were measured after the subject had fine-tuned the system, whereas in the case of the automatic control (AUTO-AR), they were the values estimated by the stereovision grasp decoder. In the SEMI conditions, we also registered if the user corrected the controller decisions (aperture size, grasp type). We assumed that the user corrected

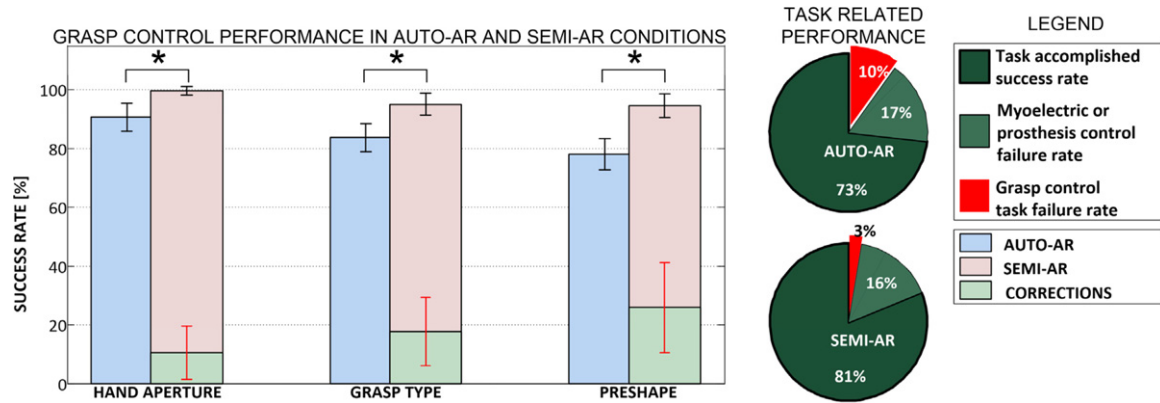


Figure 7. The grasp control and task related performance in the AUTO-AR and SEMI-AR experimental conditions. The overall rate of user corrections during the SEMI-AR is also depicted. The bars represent the mean and standard deviation of the grasp control outcome measures (GC2–4, table 2). The statistically significant differences are denoted by a star ($p < 0.05$). The pie charts depict the per cent rate for the task related performance indices (TA1–2, table 2). Note that the semi-automatic control resulted in significantly improved performance.

the preshape if he/she corrected either the grasp type or the aperture size. We used a two-sided t -test for dependent samples (repeated measures) for pair-wise comparison between the conditions. A p -value of 0.05 was selected as the threshold for statistical significance.

3. Results

Thirteen subjects performed a total of 1560 trials (13 subjects \times 6 series \times 20 objects). Thus, the stereovision grasp decoder estimated a grasp type and an aperture size 1560 times. Overall, correct estimations of size and grasp type accounted for 91% and 90% of the cases, respectively. The preshape success rate (i.e., correct grasp type and aperture size) was 84%. The average size estimation error was 0.73 ± 1.08 cm (mean \pm standard deviation).

Figure 7 summarizes the results for the first (AUTO-AR) and second (SEMI-AR) experimental conditions which included 520 (13 subjects \times 2 series \times 20 objects) and 260 trials (13 subjects \times 1 series \times 20 objects), respectively. When the user was allowed to correct for the errors (SEMI-AR), the performance improved significantly (bars in figure 7). The average improvement was 10%, 11% and 16% for GC2 (hand aperture success rate), GC3 (grasp type success rate) and GC4 (preshape success rate), respectively. The resulting performances were close to 100%: $99 \pm 5\%$, $94 \pm 9\%$ and $94 \pm 5\%$ for GC2, GC3 and GC4, respectively. The subjects intervened in 25% of the cases either by adjusting the hand aperture size or grasp type. A representative example of the correction of the hand aperture is shown in figure 8. The task accomplishment outcome measures (TA1–2) are shown in the pie charts in figure 7. During the AUTO-AR condition the task was successfully completed in 73% of trials. Successful completion of the task increased significantly to 81% in the SEMI-AR condition ($p = 0.041$). This improvement was due to a reduction of the *grasp control task failure rate* (TA2 in table 2), which dropped from 11% to only 3%. Therefore, when given the opportunity to correct and/or fine-tune the decisions of the decoder, the users were able to drive the prosthesis into

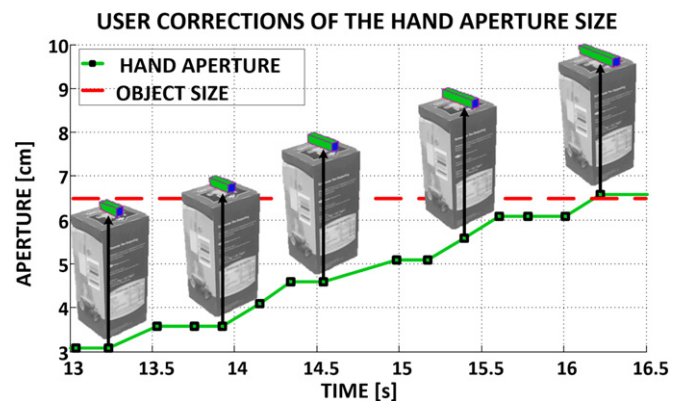


Figure 8. The sequence of user corrections of the hand aperture size. The subject relied on the AR feedback to implement the corrections while the hand was outside his/her field of view. The size of the AR box (green box) was initially smaller with respect to the size of the target object and the subject gradually opened the hand through a sequence of discrete steps (black points along the plot) until the two sizes approximately matched.

postures that yielded more stable grasps in the next phase, and thus were able to significantly improve the task completion rate (TA1). Furthermore, the average time the user spent for the task accomplishment (TTA) in the AUTO-AR scenario was 2.77 s in contrast to 3.47 s in the SEMI-AR (with no statistically significant difference, $p = 0.065$). Therefore, we can conclude that in the SEMI-AR mode the user achieved significantly better performance compared to the AUTO-AR without significantly increasing the average time needed to accomplish the task. Thus, the semi-automatic control mode of the presented system is preferred, as it offers clear performance advantages, with respect to the fully automatic control.

The results for the SEMI-AR-RE and SEMI-VIS-RE experimental conditions are presented in figure 9. In these conditions, random errors were added to the estimated aperture size. Consequently, the rate of user corrections for grasp

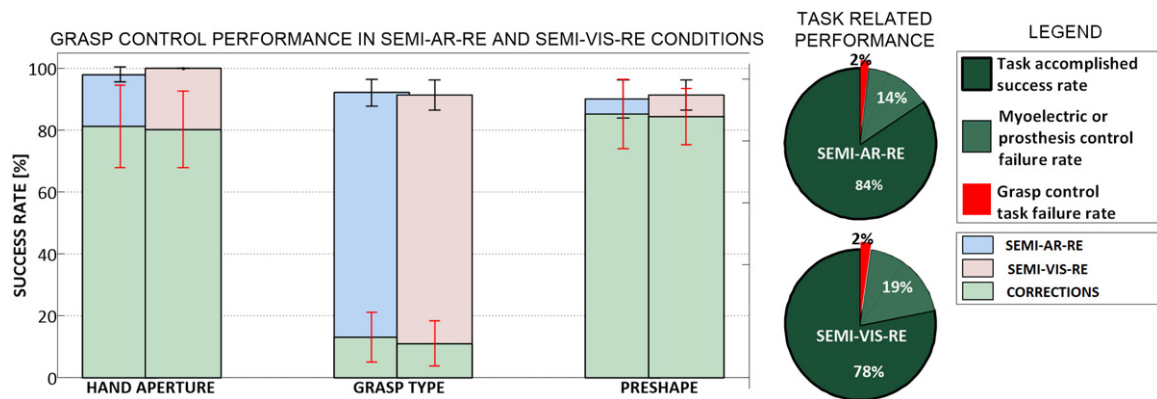


Figure 9. The grasp control and task related performance in the SEMI-AR-RE and SEMI-VIS-RE experimental conditions. The overall rate of user corrections during both scenarios is also depicted. The bars represent the mean and standard deviation of the grasp control outcome measures (GC2–4, table 2). The pie charts depict the per cent rate for the task related performance indices (TA1–2, table 2). Note that the performance measures and rate of user corrections were very similar in the two conditions.

size and the time needed to accomplish the task (5.2 s in SEMI-AR-RE and 5.9 s in SEMI-VIS-RE) increased significantly compared to the SEMI-AR condition, whereas the rate of grasp type corrections did not statistically differ. The outcome measures achieved with SEMI-AR-RE (figure 9) did not statistically differ from those achieved with SEMI-AR (figure 7); only GC4 (preshape success rate) was statistically different ($p = 0.041$) and somewhat lower in the SEMI-AR-RE condition (90% versus 94%). Therefore, although the subject was challenged by the poor performance of the automatic controller in SEMI-AR-RE, the overall performance was almost unaffected due to the possibility of manually adjusting the initial decisions of the stereovision grasp decoder. SEMI-AR-RE and SEMI-VIS-RE also resulted in a similar performance with no statistically significant differences ($p = 0.504$, $p = 0.407$, $p = 0.727$, $p = 0.209$, and $p = 0.977$ for GC2, GC3, GC4 and TA1, TA2 respectively). There was also no statistically significant difference in the time needed to accomplish the task (TTA) in these two conditions ($p = 0.492$). The fact that all performance measures were similar in both SEMI-VIS-RE and SEMI-AR-RE demonstrates that the subjects successfully utilized the AR feedback to correct the mistakes of the automatic controller. Therefore, the AR feedback is indeed a feasible medium to provide information to the user, i.e., in this specific case artificial-proprioception.

4. Discussion

As previously stated in introduction, the main contribution of this paper is a novel system that integrates volitional, biological and automatic, artificial control in order to decrease the burden on the user. This was achieved by fusing the information from two sources, i.e., artificial vision and myoelectric signals. In addition, the control loop was closed through the application

of AR feedback, which is proposed as a method to further promote this integration by communicating to the user not only the prosthesis state but also internal information of the controller. The specific context (practical human–machine interaction) of the proposed method is directly reflected in the design decisions and selection of the system components. For example, in order to be robust and general, the vision module does not assume a predefined database of concrete objects. Also, the AR glasses are a wearable component which is likely to become even more ergonomic and convenient for the practical application due to the fast development of this technology (Google glasses [46]). We assessed the feasibility of this system by implementing a prototype that will be developed further in the future. The test was done in healthy subjects using an artificial hand prosthesis as a convenient experimental platform, but the proposed concept could be generalized to many assistive technologies for the restoration of reaching and grasping.

4.1. Experimental evaluation

The overall results for the performance of the stereovision grasp decoder (1560 trials in total) demonstrate that stereovision can be used to estimate the grasp relevant properties in a set of target objects with different shapes and sizes. The experimental task showed that after only a few minutes of training, the subjects were able to operate the system successfully and use the fully automatic control of grasping.

The autonomous operation was not flawless but the subjects were able to correct most of the errors by exploiting the AR feedback to fine-tune the decisions of the stereovision grasp decoder (semi-automatic control). The subjects successfully understood the meaning of the AR feedback, perceiving correctly the information about the grasp type and size. Also, subjects understood the bidirectional control loop, the concept of control sharing and they operated the system smoothly through different phases comprising the full control cycle (i.e., TTA was only few seconds).

They successfully used the AR feedback information about the prosthesis state as well as controller decisions to make appropriate corrective actions when needed, bringing the control of grasping very close to an ideal performance ($\sim 100\%$). Thus, all of this demonstrates the feasibility of the overall conceptual solution and its components, such as stereovision grasp decoder and AR bidirectional communication interface. In addition, the rate of failure in the task accomplishment due to incorrect grasp control (TA2) was consistently very small ($\sim 2\text{--}3\%$) in all the SEMI-AR conditions. Importantly, the transition from automatic to semi-automatic control was almost effortless and consisted in the introduction of a few additional simple myoelectric commands to adjust the decisions of the stereovision grasp decoder. The closed-loop system demonstrated very good robustness when the user was challenged by the introduction of the intentional errors into the automatic control (SEMI-AR-RE). The errors were present in every trial and the error size was such that in most of the cases it would have been impossible to accomplish the grasp without a correction. However, this did not influence the overall performance.

Finally, the feasibility of the AR feedback was further confirmed by the last test (SEMI-AR-RE versus SEMI-VIS-RE) which demonstrated that the subjects were able to exploit the novel AR feedback as easily as the normal visual feedback. The subjects perceived the virtual object embedded into the real scene and were able to correctly compare its size against the size of the real objects as easily as when comparing the size of the real objects directly (hand opening versus object size). It is important to emphasize that the scenarios SEMI-AR-RE and SEMI-VIS-RE do not necessarily reflect the expected use of the AR and normal visual feedback in the potential real application. It is likely that trained subjects will learn to rely more on the feed-forward control and use feedback only when necessary.

When analyzing the task completion success rate across the experimental scenarios it is notable that the behavior of the conglomerate system was also affected by failures unrelated to the automatic/semi-automatic control of grasping, but to the myoelectric control and prosthesis design (myoelectric/prosthesis control failure rate (figures 7 and 9)). This depended on several factors: (1) individual ability of the subject to generate and control EMG signals as well as dexterity in prosthesis handling, (2) intrinsic properties of the prosthetic device and its mounting system and (3) a simple method to process the EMG signals. Therefore, these errors could be minimized by providing the prospective users with more training so that they can generate more consistent muscle activations while manipulating the prosthesis, by improving the prosthesis mechanical design (e.g., silicone coating over the fingers to prevent object slipping), and by implementing more robust methods for myoelectric control. These errors are general issues common to any prosthetic system and, as explained before, they do not reflect the performance of the core feature of the novel approach (semi-automatic control of grasping).

4.2. Computer vision and automatic control of grasping

Computer vision analysis was based on the standard state-of-the-art methods. The performance in the automatic control of grasping achieved during the experimental sessions indicates that the computer vision module was relatively robust and reliable. The robustness is further illustrated in figure 10, which demonstrates the system operation in a rich, cluttered environment. The figure depicts several scenes in which the subject successfully selected and grasped four target objects. It is worth noting that in both cases the scene included many other objects in addition to the target. Nevertheless, the system correctly located and segmented the target object as shown by the semi-transparent green overlay representing the pixels allocated to the object (figure 10, left). The segmentation, modeling, size estimation and grasp type selection (figure 10, right) were all correct despite the richness of the scene and the fact that the target object was partially occluded by the object(s) placed in front of it.

The analysis of the scene by using stereovision is much more powerful and robust compared to the use of a single camera [35], but still suffers from well-known limitations like pixel correspondence problems [51]. This was also registered in our experiments. During the object targeting phase, the computer vision module sometimes failed to segment out the target object or there was a ‘spillover’ from the area of the target into the neighboring objects or the background. However, this was quite a rare event ($<3\%$) and it did not significantly influence the performance due to the following countermeasures: (1) the ‘snapshot and analysis’ was repeated continuously during targeting and an occasional failure would be most often corrected in the very next cycle; (2) the users could rely on the AR feedback to assess the quality of segmentation and avoid triggering when large errors were evident.

The system developed is an illustrative example of how an artificial controller can be enriched with an additional, non-conventional information source (stereo camera pair) and high level processing (cognitive like reasoning) to achieve fully automatic control of the functions that are conventionally the responsibility of the user (e.g., hand preshaping). In this scheme, the user is able to ‘release’ predefined ‘motor programs’ performing relatively complex functions, instead of continuously monitoring the task. This substantially simplifies the myoelectric interface, which only needs to implement a simple triggering mechanism, thereby reducing the burden on the user. This can be advantageous especially in the case of modern dexterous prosthetic hands and/or full arm prostheses. The presented control concept scales smoothly with the system complexity. In fact, the more complex the system is, the higher and more obvious will be a discrepancy between the necessary user efforts and required system functionality. For example, in the case of an entire upper limb prosthesis, the stereovision control could be used both to preshape the hand and to navigate the arm to reach and grasp the selected target object (stereovision servoing [52]). The complex ‘preshape and reach program’ could be triggered via a simple myoelectric command.

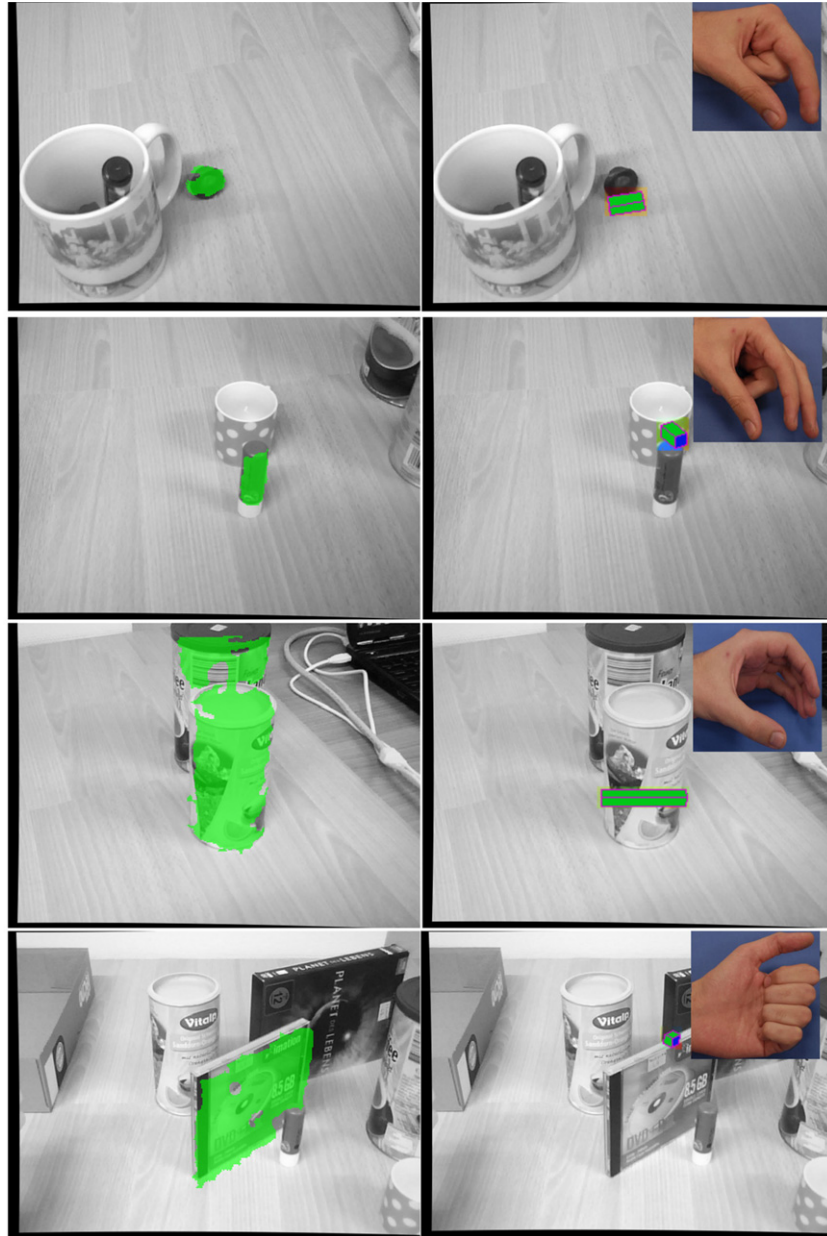


Figure 10. An example of system operation in a cluttered environment. The subject targeted several different objects (left side) placed within a realistic scene including several other colorful and textured objects and a background. The stereovision grasp decoder successfully segmented out the targets and selected an appropriate grasp type (grasp type icons) and size (the AR box with size similar to the target object) for each of them (right side).

4.3. Augmented reality feedback and semi-automatic control

The idea to close the loop ‘through’ the user by providing feedback about the state of the prosthesis is not novel [53]. However, the current work proposes a fundamentally novel interface to accomplish this goal. Compared to the ‘classical’ methods of electrical or direct mechanical stimulation, the most important advantage of the AR feedback is that it can exploit the high bandwidth and flexibility that is available within the visual communication channel. This is reflected both in the sheer volume of information that can be communicated to the subject as well as in the form in which

that information can be presented. In the current work, the AR feedback was projected into the scene, next to the target object and a virtual box was used to implement ‘artificial proprioception’, i.e., the size of the box was proportional to the actual aperture size of the hand. As stated before, the system operation and experimental paradigm were not specifically designed to test the usefulness of the AR feedback and/or the potential advantages of the AR versus normal visual feedback. Instead, the aim was only to test the general feasibility since the AR interface has been used in this context for the first time, i.e., testing if the AR interface can be used as a component for implementing user corrections within a novel

control system. However, even this first basic implementation demonstrates some possible benefits. Namely, as explained in the experimental protocol, in SEMI-AR and SEMI-AR-RE scenarios, the users operated the hand (i.e., correcting grasp type and size) without seeing it, i.e., the hand was outside the field of view and the subjects were looking at the object to be grasped. Note that this corresponds to the way grasping is performed in daily life by able-bodied subjects. Apart from promoting the normal use of the hand, artificial proprioception through AR can have some obvious advantages in certain situations, e.g., when the hand is partially occluded by some other object or is not completely visible from the user perspective, which is often the case during grasping (especially in cluttered environments). Moreover, the potential utility of AR feedback is even more evident when taking into account that the feedback could be much more sophisticated; for example, a full graphical model of the hand/prosthesis could be used to communicate complete and very detailed information in an intuitive way (e.g., position, contact and force of each individual finger). The flexibility and bandwidth of AR feedback were exploited in some other fields (e.g., robot-assisted surgeries [54]), but not in prosthetics which is a context with unique requirements. How to best utilize the potential of AR feedback in daily life prosthetic or clinical use is a very interesting and relevant question that will be addressed in the studies to follow.

To minimize the interference of AR feedback with the ongoing visual or other cognitive tasks, AR feedback could be simplified (e.g., a 2D object) and moved within the peripheral part of the visual field. The optimal form and positioning of AR feedback is an important question that will be addressed in future studies. It is also possible to combine several representations to best reconcile the needs of different tasks (e.g., unobtrusive, simple feedback most of the time and detailed feedback during specific, more demanding tasks).

A further novelty that was proposed in this work is the use of feedback not only to acknowledge to the user about the state of the device but also to allow him to monitor the decisions of the artificial controller. This gives the opportunity to the subject to supervise the automatic operation and take over control when needed, effectively implementing a bilateral communication between the user and controller. In the semi-automatic framework, the control is therefore shared between the two agents (user and controller) and the optimal integration of the two control loops (manual and automatic) will be an important question to address in future research.

4.4. Perspectives and future work

An important feature of the presented concept and corresponding implementation is a modular design. Strictly

speaking, both stereovision and AR feedback could be used on their own, independently of each other. For example, the stereovision could implement automatic control, while the corrections are performed using conventional visual feedback (as in the SEMI-VIS-RE condition). Or, the AR could be used to provide feedback about the prosthetic device that is controlled 'manually' using a classic myoelectric control. The placement of the sensor element is also flexible. In the current implementation, the cameras were embedded within the front glasses, but a miniature sensor could be placed on the sides or above the glasses, for example, into the glass frame, or even incorporated into the hand or a prosthetic socket. Truly wearable and cosmetically acceptable, mobile solutions for the AR, such as specialized glasses and even contact lenses, are being developed and some models (e.g., Google Glass project [46]) are already available in limited quantities and are expected to become widely accessible very soon. This, together with the development and availability of fast processing cores within small form factor embedded systems, can provide the necessary technological framework for the practical implementation of the presented control concept. From the technical standpoint, the developed system relies on state-of-the-art technologies that are in the focus of current research efforts and are also fast developing. Although some of the components might not be ready for immediate practical application in daily life, this might change in the very near future. Stereovision is not the only method to implement the proposed control concept. The necessary information for the control of grasping is an estimate of the 3D shape of the target object and any technology providing a similar output (i.e., 3D point cloud) could be plugged into the control scheme in a straightforward manner. For example, the active methods based on IR depth sensing derived from Microsoft Kinect are more robust than passive stereovision. In general, this technology is developing very fast, evolving towards solutions that will soon become low cost and physically small, and thereby very convenient for practical applications (e.g., Prime Sense [55], Creative Senz3D [56]). Finally, advanced scene processing [57] and grasp planning methods [58] that were investigated intensively in robotics research could be used to improve the prosthesis preshape module, but the candidate solutions could be only those approaches that can cope with the constraints of this specific application (e.g., incomplete input information, responsiveness to user decisions, *ad hoc* scene etc).

Importantly, the presented approach and its components are a rather general solution that could be applied with relatively minor adjustments in many contexts. In principle, it could be used with any assistive system that includes a multi-grasp end effector (e.g., electrical stimulator, hand exoskeleton, hand prosthesis). To port the control scheme (figure 3) to a different device, the prosthesis control module would have to be replaced with a module providing a low level interface to that specific system. All the other components could remain unchanged or, in the worst case, they would have to be somewhat adapted to support the new application (e.g., available grasps). Alternatively, only a part of the functionality

could be ported. For example, stereovision grasp decoder and myoelectric control could be used for grasp type and size selection in EMG-triggered FES. Similarly, AR glasses could be used not for control as in this study, but only to provide AR feedback, supplementing the impaired sensations in patients suffering from a neurological condition (e.g., stroke). As explained below, stereovision can be extended to the control of orientation, and in this case, the scope of potential applications is even larger: a rehabilitation robot for assistance in reaching, full arm prosthesis, a hybrid system combining a reaching robot and an FES device for grasping. Finally, the presented control system could indeed be used in daily life for prosthetic or orthotic applications, which is similar to the scenario presented in the current study, but also only during a limited time period, for therapy (e.g., functional electrical therapy, robotic rehabilitation) or training (e.g., simplifying myoelectric training for a complex prosthesis).

The presented system is an example of the sensor fusion approach to the control of prosthetic devices [34]. In the current system, control was implemented by combining visual information and myoelectric commands. This approach could be extended further. The 3D information about the scene and the target object could be integrated with information about the pose of the prosthesis provided by inertial measurement sensors placed on the device, resulting in an adaptive system that can reconfigure automatically depending on the side and angle of 'attack'. For example, the hand aperture could readjust depending on the side the user is approaching the object or in the case the hand is equipped with wrist rotator and/or flexion/extension, the system can control the hand orientation. The future steps will be system improvement, in terms of more functions and practicality, followed by an evaluation in the specific context and on the actual target population, i.e., amputees or patients, depending on the selected application. The goal of these tests will be to assess the usability, acceptance and finally the efficiency of the proposed interface.

We believe that there are many potential benefits of using the AR interface versus normal visual feedback or tactile stimulation, and investigating this will be the topic of the studies to follow. For example, as a separate study, we plan to perform an in-depth analysis of the performance and possible applications of proposed AR feedback concepts. Specifically, one of the goals would be to compare the regular open-loop (visual inspection) and closed-loop (AR feedback) prosthesis control in more complex real-life scenarios (e.g., cluttered environment with occlusions, adjusting the hand while reaching for the object). The current AR interface could also be easily extended to feedback additional signals. Specifically, the AR channel could be used for the closed-loop

control of the grasping force or prosthesis orientation. For example, the grasping force could be visualized as a vertical bar in the peripheral visual field with the height of the bar being proportional to the force amplitude. These developments are also something that we plan to address in future studies, as we believe they are the key for unlocking the full potential of the AR feedback interface.

We are well aware that there are drawbacks that could jeopardize some of the potential applications. For example, when applied in prosthetic or orthotic scenarios, the system could provide simple and effective control of a complex device, but at the expense of adding additional components. In the current version, AR glasses with stereo camera pair have to be worn by the subject. For the user of a prosthesis, this could be an additional nuisance that could further compromise the already sensitive process of device acceptance. However, as always, this ultimately depends on the cost benefit ratio: if the user gets a complex multi-degree of freedom prosthesis that can react to a simple command by automatically reaching and grasping for the desired object, we could expect that he/she would likely be much more eager to wear an additional component. In fact, the control scheme proposed here was filed as a joint patent [59] with our industrial partner (Otto Bock Healthcare GmbH, Vienna, AT), one of the leading manufacturers of prosthetic equipment. Importantly, there are some recent technological developments that could significantly improve the system appearance, thereby overcoming the present drawbacks. Finally, this issue is less likely to be a serious obstacle when considering the potential short term system application (therapy or training). There are studies reporting the application of virtual reality equipment in patients with paralysis [60] as well as in amputees [61]. In conclusion, this study presented a set of novel methods and demonstrated their feasibility using a first developed prototype. The next steps are to further improve the system (ongoing work) and then to benchmark the novel approach against the current state-of-the-art prosthetic systems (semi-automatic versus manual myoelectric control, normal visual and/or tactile feedback versus AR feedback).

Acknowledgments

We acknowledge financial support by the German Ministry for Education and Research (BMBF) via the Bernstein Focus Neurotechnology (BFNT) Göttingen (grant 1GQ0810) (DF) and the Italian Ministry of Education University and Research, under the FIRB-2010 MY-HAND Project (RBFR10VCLD).

Appendix. Control flow in detail

Table A1. Control flow.

I. Object targeting phase

The hand opens

WHILE TRUE

Analyze the scene and look for the object closest to the central point of the image.

IF the object detected THEN highlight the object by placing a pink overlay (AR feedback).

IF extensor EMG low/high detected THEN:

- 1) Highlight the object by placing a green overlay (AR feedback)
- 2) Determine a geometrical model of the object and its properties (shape and size)
- 3) Select the grasp type and size
- 4) GOTO PHASE II.

END

II. Hand preshaping phase

The hand preshapes according to the selected grasp type (palmar, lateral, bidigit pinch and tridigit pinch) and size (continuous estimation).

Indicate to the user the selected grasp type for 3 s (AR feedback).

WHILE TRUE

A virtual 3D box object corresponding to the current prosthesis aperture is augmented into the real scene (AR feedback).

IF extensor EMG high detected THEN manual correction (restart):

- 1) Stop feedback
- 2) Open hand
- 3) GOTO PHASE I

ELSE IF extensor EMG low detected THEN manual correction (aperture):

- 1) Increase the prosthesis aperture by 0.5 cm
- 2) Update the hand aperture AR feedback

ELSE IF flexor EMG low detected THEN manual correction (aperture):

- 1) Decrease the prosthesis aperture by 0.5 cm
- 2) Update the hand aperture AR feedback

ELSE IF flexor EMG high detected THEN trigger hand closing:

- 1) Stop feedback.
- 2) Start closing the prosthesis.
- 3) GOTO PHASE III.

END

END

III. Manipulation phase

The hand closes.

WHILE TRUE

IF extensor EMG low/high detected THEN trigger hand opening:

- 1) Start opening the prosthesis.
- 2) GOTO PHASE I.

END

END

References

- [1] MacKenzie C and Iberall T 2010 *The Grasping Hand* (Amsterdam: Elsevier)
- [2] Bootsma R J and van Wieringen P C W 1992 Spatio-temporal organisation of natural prehension *Hum. Mov. Sci.* **11** 205–15
- [3] Connolly J D and Goodale M a 1999 The role of visual feedback of hand position in the control of manual prehension *Exp. Brain Res.* **125** 281–6
- [4] Popović M B 2003 Control of neural prostheses for grasping and reaching *Med. Eng. Phys.* **25** 41–50
- [5] Balasubramanian S, Klein J and Burdet E 2010 Robot-assisted rehabilitation of hand function *Curr. Opin. Neurol.* **23** 661–70
- [6] Cipriani C, Controzzi M and Carrozza M C 2011 The SmartHand transradial prosthesis *J. Neuroeng. Rehabil.* **8** 29
- [7] Heo P, Gu G M, Lee S, Rhee K and Kim J 2012 Current hand exoskeleton technologies for rehabilitation and assistive engineering *Int. J. Precis. Eng. Manuf.* **13** 807–24
- [8] Touch Bionics iLimb 2014 www.touchbionics.com/products/active-prostheses/
- [9] Bebionic Hand 2014 http://bebionic.com/the_hand
- [10] Parker P, Englehart K and Hudgins B 2006 Myoelectric signal processing for control of powered limb prostheses *J. Electromyogr. Kinesiol.* **16** 541–8
- [11] Xiao R and Ding L 2013 Evaluation of EEG features in decoding individual finger movements from one hand *Comput. Math. Methods Med.* **2013** 243257
- [12] Pistohl T, Schulze-Bonhage A, Aertsen A, Mehring C and Ball T 2012 Decoding natural grasp types from human ECoG *NeuroImage* **59** 248–60
- [13] Carrozza M C, Persichetti A, Laschi C, Vecchi F, Lazzarini R, Vacalebri P and Dario P 2007 A wearable biomechatronic

- interface for controlling robots with voluntary foot movements *IEEE/ASME Trans. Mechatronics* **12** 1–11
- [14] Johansen D, Popović D B, Struijk L N S A, Sebelius F and Jensen S 2012 A novel hand prosthesis control scheme implementing a tongue control system *Int. J. Eng. Manuf.* **2** 14
- [15] Controzzi M, Cipriani C, Popovic D B and Carrozza M C 2013 Controlling hand-assistive devices: utilizing electrooculography as a substitute for vision *IEEE Robot. Autom. Mag.* **20** 40–52
- [16] Asghari Oskoei M and Hu H 2007 Myoelectric control systems—a survey *Biomed. Signal Process. Control* **2** 275–94
- [17] Peerdeman B et al 2011 Myoelectric forearm prostheses: state of the art from a user-centered perspective *J. Rehab. Res. Develop.* **48** 719–37
- [18] Tenore F V G, Ramos A, Fahmy A, Acharya S, Etienne-Cummings R and Thakor N V 2009 Decoding of individuated finger movements using surface electromyography *IEEE Trans. Biomed. Eng.* **56** 1427–34
- [19] Cipriani C, Antfolk C, Controzzi M, Lundborg G, Rosen B, Carrozza M C and Sebelius F 2011 Online myoelectric control of a dexterous hand prosthesis by transradial amputees *IEEE Trans. Neural Syst. Rehabil. Eng.* **19** 260–70
- [20] Micera S et al 2011 Decoding of grasping information from neural signals recorded using peripheral intrafascicular interfaces *J. Neuroeng. Rehabil.* **8** 53
- [21] Baker J J, Scheme E, Englehart K, Hutchinson D T and Greger B 2010 Continuous detection and decoding of dexterous finger flexions with implantable myoelectric sensors *IEEE Trans. Neural Syst. Rehabil. Eng.* **18** 424–32
- [22] Szeto A Y and Saunders F A 1982 Electrocutaneous stimulation for sensory communication in rehabilitation engineering *IEEE Trans. Biomed. Eng.* **29** 300–8
- [23] Kaczmarek K A, Webster J G, Bach-y-Rita P and Tompkins W J 1991 Electrotactile and vibrotactile displays for sensory substitution systems *IEEE Trans. Biomed. Eng.* **38** 1–16
- [24] Patterson P E and Katz J A 1992 Design and evaluation of a sensory feedback system that provides grasping pressure in a myoelectric hand *J. Rehabil. Res. Dev.* **29** 1–8
- [25] Meek S G, Jacobsen S C and Goulding P P 1989 Extended physiologic tactation: design and evaluation of a proportional force feedback system *J. Rehabil. Res. Dev.* **26** 53–62
- [26] Scott R N, Brittain R H, Caldwell R R, Cameron A B and Dunfield V A 1980 Sensory-feedback system compatible with myoelectric control *Med. Biol. Eng. Comput.* **18** 65–69
- [27] Pylatiuk C, Kargov A and Schulz S 2006 Design and evaluation of a low-cost force feedback system for myoelectric prosthetic hands *J. Prosthet. Orthot.* **18** 57–61
- [28] Cipriani C, D'Alonzo M and Carrozza M C 2012 A miniature vibrotactile sensory substitution device for multifingered hand prosthetics *IEEE Trans. Biomed. Eng.* **59** 400–8
- [29] D'Alonzo M, Dosen S, Cipriani C and Farina D 2013 HyVE: hybrid vibro-electrotactile stimulation for sensory feedback and substitution in rehabilitation *IEEE Trans. Neural Syst. Rehabil. Eng.* **22** 290–301
- [30] Dhillon G S and Horch K W 2005 Direct neural sensory feedback and control of a prosthetic arm *IEEE Trans. Neural Syst. Rehabil. Eng.* **13** 468–72
- [31] Karnati N, Kent B A and Engeberg E D 2013 Bioinspired sinusoidal finger joint synergies for a dexterous robotic hand to screw and unscrew objects with different diameters *IEEE/ASME Trans. Mechatronics* **18** 612–23
- [32] Malhotra M, Rombokas E, Theodorou E, Todorov E and Matsuoka Y 2012 Reduced dimensionality control for the ACT hand *IEEE Int. Conf. on Robotics and Automation* pp 5117–22
- [33] Matrone G C, Cipriani C, Carrozza M C and Magenes G 2012 Real-time myoelectric control of a multi-fingered hand prosthesis using principal components analysis *J. Neuroeng. Rehabil.* **9** 40
- [34] Dosen S, Muller K-R and Farina D 2012 Myoelectric control of artificial limbs—is there a need to change focus? *IEEE Signal Process. Mag.* **29** 150–2
- [35] Dosen S, Cipriani C, Kostić M, Controzzi M, Carrozza M C and Popović D B 2010 Cognitive vision system for control of dexterous prosthetic hands: experimental evaluation *J. Neuroeng. Rehabil.* **7** 42
- [36] Tomovic R and Boni G 1962 An adaptive artificial hand *IRE Trans. Autom. Control* **7** 3–10
- [37] Tomovic R, Bekey G and Karplus W 1987 A strategy for grasp synthesis with multifingered robot hands *Proc. IEEE Int. Conf. on Robotics and Automation* vol 4 pp 83–89
- [38] Chappell P H, Nightingale J M, Kyberd P J and Barkhordar M 1987 Control of a single degree of freedom artificial hand *J. Biomed. Eng.* **9** 273–7
- [39] Otto Bock Sensor Hand Speed 2014 www.ottobock.de/cps/rde/xchg/ob_com_en/hs.xsl/3652.html
- [40] Engeberg E D and Meek S G 2013 Adaptive sliding mode control for prosthetic hands to simultaneously prevent slip and minimize deformation of grasped objects *IEEE/ASME Trans. Mechatronics* **18** 376–85
- [41] Wettels N, Parnandi A R, Loeb G E and Sukhatme G S 2009 Grip control using biomimetic tactile sensing systems *IEEE/ASME Trans. Mechatronics* **14** 718–23
- [42] Light C M, Chappell P H, Hudgins B and Engelhart K 2002 Intelligent multifunction myoelectric control of hand prostheses *J. Med. Eng. Technol.* **26** 139–46
- [43] Cipriani C, Zacccone F, Micera S and Carrozza M C 2008 On the shared control of an EMG-controlled prosthetic hand: analysis of user–prosthesis interaction *IEEE Trans. Robot.* **24** 170–84
- [44] Corbett E A, Kording K P and Perreault E J 2013 Real-time evaluation of a noninvasive neuroprosthetic interface for control of reach *IEEE Trans. Neural Syst. Rehabil. Eng.* **21** 674–83
- [45] Kragic D, Miller A T and Allen P K 2001 Real-time tracking meets online grasp planning *ICRA'01: Proc. IEEE Int. Conf. on Robotics and Automation* vol 3 pp 2460–5 (Cat. No.01CH37164)
- [46] Google Glass Project 2014 www.google.com/glass/start/
- [47] Geiger A, Roser M and Urtasun R 2010 *Computer Vision—ACCV (Lecture Notes in Computer Sciences* vol 6492) (Berlin: Springer)
- [48] Bouguet J-Y 2013 *Camera Calibration Toolbox for Matlab* www.vision.caltech.edu/bouguetj/calib_doc/
- [49] Fischler M A and Bolles R C 1981 Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography *Commun. ACM* **24** 381–95
- [50] Cutkosky M R 1989 On grasp choice, grasp models, and the design of hands for manufacturing tasks *IEEE Trans. Robot. Autom.* **5** 269–79
- [51] Siebert J P and Cyganek B 2009 *An Introduction to 3D Computer Vision Techniques and Algorithms* (New York: Wiley) p 504
- [52] Hutchinson S, Hager G D and Corke P I 1996 A tutorial on visual servo control *IEEE Trans. Robot. Autom.* **12** 651–70

- [53] Childress D S 1980 Closed-loop control in prosthetic systems: historical perspective *Ann. Biomed. Eng.* **8** 293–303
- [54] The da Vinci® Surgical System 2014 www.davincisurgery.com/da-vinci-surgery/da-vinci-surgical-system/
- [55] Prime Sense Sensors 2014 www.primesense.com/
- [56] Creative Sens3D 2014 <http://de.creative.com/p/web-cameras/creative-senz3d>
- [57] PCL 2014 <http://pointclouds.org/>
- [58] Ciocarlie M T and Allen P K 2009 Hand posture subspaces for dexterous robotic grasping *Int. J. Rob. Res.* **28** 851–67
- [59] Dosen S, Markovic M, Popović D B, Farina D and Graimann B Sensor fusion for control of upper limb protheses: integration of myoelectric control with stereovision, augmented reality and inertial sensing *Patent number* EP 13171671.42013
- [60] Subramanian S, Knaut L A, Beaudoin C, McFadyen B J, Feldman A G and Levin M F 2007 Virtual reality environments for post-stroke arm rehabilitation *J. Neuroeng. Rehabil.* **4** 20
- [61] Murray C D, Pettifer S, Howard T, Patchick E L, Caillette F, Kulkarni J and Bamford C 2007 The treatment of phantom limb pain using immersive virtual reality: three case studies *Disabil. Rehabil.* **29** 1465–9