Available online at www.sciencedirect.com
SciVerse ScienceDirect
Int. J. Human-Computer Studies 71 (2013) 590–607
International Journal of Human-Computer Studies
www.elsevier.com/locate/ijhcs

Automatic recognition of object size and shape via user-dependent measurements of the grasping hand

Radu-Daniel Vatavu*, Ionuţ Alexandru Zaiţi

University Stefan cel Mare of Suceava, str. Universitatii nr. 13, 720229 Suceava, Romania

Received 4 July 2011; received in revised form 13 August 2012; accepted 3 January 2013
Communicated by E. Motta
Available online 24 January 2013

Abstract

An investigation is conducted on the feasibility of using the posture of the hand during prehension in order to identify geometric properties of grasped objects such as size and shape. A recent study of Paulson et al. (2011) already demonstrated the successful use of hand posture for discriminating between several actions in an office setting. Inspired by their approach and following closely the results in motor planning and control from psychology (Makenzie and Iberall, 1994), we adopt a more cautious and punctilious approach in order to understand the opportunities that hand posture brings for recognizing properties of target objects. We present results from an experiment designed in order to investigate recognition of object properties during grasping in two different conditions: object translation (involving firm grasps) and object exploration (which includes a large variety of different hand and finger configurations). We show that object size and shape can be recognized with up to 98% accuracy during translation and up to 95% and 91% accuracies during exploration by employing user-dependent training. In contrast, experiments show less accuracy (up to 60%) for user-independent training for all tested classification techniques. We also point out the variability of individual grasping postures resulted during object exploration and the need for using classifiers trained with a large set of examples. The results of this work can benefit psychologists and researchers interested in human studies and motor control by providing more insights on grasping measurements, pattern recognition practitioners by reporting recognition results of new algorithms, as well as designers of interactive systems that work on gesture-based interfaces by providing them with design guidelines issued from our experiment.

Keywords: Hand posture; Grasping; Prehension; Object recognition; Experiment; Shape recognition; Size recognition; Measurements; Data glove; Gestures

1. Introduction

The human hand represents a remarkable instrument for grasping and manipulating objects but also for extracting useful information such as object weight, size, orientation, surface texture, and thermal properties. Hands serve therefore both executive and perceptive functions synchronously while they are employed to explore as well as to transform and change the environment. In MacKenzie and Iberall's own words: *we use our hands as general purpose devices, to pick up objects, to point, to climb, to play musical instruments, to draw and sculpt, to communicate, to touch and feel, and to explore the world* (Makenzie and Iberall, 1994, p. 4). The understanding of the hand functions and the inner workings of the fine mechanisms for planning and controlling movements at both muscular and neural (central nervous system) levels still represent a very active field of study for psychologists interested in motor control (Jones and Lederman, 2006; Makenzie and Iberall, 1994; Wing et al., 1996). However, MacKenzie and Iberall's description of hands as instruments can prove to be extremely interesting to human–computer interaction: the hand becomes a specialized device that extracts information from the objects users are interacting with. This can be described as *extrinsic-oriented exploration* of the environment

*Corresponding author. Tel./fax: +40 230 524801.
E-mail addresses: raduvro@yahoo.com, vatavu@eed.usv.ro (R.-D. Vatavu), ionutzaiti@yahoo.com (I.A. Zaiţi).
URL: http://www.eed.usv.ro/˜vatavu (R.-D. Vatavu).

1071-5819/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ijhcs.2013.01.002

in opposition to *intrinsic-oriented exploration* where objects share identification information explicitly through various technologies such as RFID (Kim et al., 2007; Tanenbaum et al., 2011; Ziegler and Urbas, 2011) and Bluetooth (Farella et al., 2008; Siegemund and Flörkemeier, 2003). As the exploration is also voluntary and accomplished using the human hand, we can further refer to it as *extrinsic-proprioceptive exploration* as the hand posture is the only available source of data from which information about target objects is inferred.

There are two main motivations in HCI for collecting and using hand posture information during object grasping and exploration. One of them is concerned with recognizing hand postures and designing proper interaction metaphors in order to build natural gesture-based interfaces. The goal in this case is to provide natural and intuitive ways for users to interact with computing systems by leveraging the information richness delivered by the hand pose (Baudel and Beaudouin-Lafon, 1993; Erol et al., 2007; Wachs et al., 2011). The other motivation is represented by recognizing activity patterns for context-aware applications, leading HCI developments towards ubiquitous computing. Within this direction, Paulson et al. (2011) showed that various activities in an office such as dialing a number, holding a mug, typing at the keyboard, or handling the mouse can be recognized using hand posture information solely.

However, the application domain can be much extended and many advantages and interactive opportunities can be envisaged regarding the information obtained while the hand is grasping ambient objects. First, the hand as a measuring instrument relieves the costly need for embedding identification technology into practically every ambient object as would be the case for RFID tags (Ilie-Zudor et al., 2011). This is especially important when such tag installations become impractical for some scenarios in terms of cost, performance, and reliability (e.g., identification problems do occur cause of tag orientation, material type, and reader collision), and social acceptance of RFID technology hindered by ethical concerns (Want, 2006). Second, knowledge of how objects are manipulated can be exploited for enhancing existing interactions with everyday objects: a firm grip on the phone ends the current call; a firm grip on the door knob locks the office door on the way out while a gentle grasp simply closes it without locking; grasping the remote turns on the TV, etc. Video games can also benefit of enriched interaction experiences in the way that players can grab any object from the real world and use it inside the game. For example, grasping a simple stick informs a sports game that the user is holding the baseball bat; a toy pistol can be used in an action game that senses the hand in the ‘‘trigger’’ posture; grasping a ball can make the game character grab another snow ball and throw it at the opponent in a winter game. Traditional learning games for children such as wooden bricks that encourage hand motor development can become highly interactive whilst the hand posture is retrieved and analyzed: virtual guiding tutors that know when the child grasped the object with the right size and shape or automated monitoring of the child's progress. All such interaction opportunities with everyday objects become viable once information can be inferred about the grasped object using the hand alone.

The idea of using hand measurements in order to identify properties of target objects has been investigated before in various forms and for various purposes outside the HCI community. For example, an early study of Buchholz et al. (1992) was concerned with proposing a model for the human hand in order to noninvasively predict postures as the hand grasps different sizes of objects. The model served to predict and evaluate the prehensile capacities of the hand when grasping ellipsoidal objects (Buchholz and Armstrong, 1992) in order to provide assistance for designers of tool handles. In the psychology line of motor control work, Santello and Soechting (1998) showed that it is possible to discriminate between concave and convex objects using the relative flexure between the index and little fingers of the grasping hand.

Such previous works suggest the potential of using the hand in order to extract information about the environment, and, more precisely, about the objects users are interacting with. The findings, although dispersed in specific contexts and research communities, suggest interesting opportunities for human–computer interaction. However, in order for the community to benefit of such a technique, solid evidence and analysis must be provided for researchers and practitioners to rely on when designing and evaluating their applications. Inspired by the well-grounded results in psychology concerning the grasping hand (Jones and Lederman, 2006; Makenzie and Iberall, 1994; Wing et al., 1996) and following the results of Paulson et al. (2011), we explore the feasibility of employing hand posture to automatically extract object properties. However, in opposition to Paulson et al. (2011) that only show that a number of distinct office activities can be recognized, we adopt a more thorough procedure. By following closely the results from motor control theory concerning the act of prehension and grasping target objects, we designed an experiment for determining whether *size*, *shape*, and *size and shape* together can be reliably identified for objects with basic geometries. The analysis was carried out in two different scenarios: object translation, for which stable and firm hand grasps are used, and object exploration, for which a large variety of different hand postures and finger configurations are employed. The exploration scenario was specifically introduced in order to evaluate the technique independently of the intended use of the object and therefore to better understand its feasibility.

The contributions of our work include:

i. We show that object size and shape can be robustly inferred from measurements on the grasping hand with up to 98% accuracy during object translation and 95% and 91% accuracies during object exploration in the user-dependent training scenario.
ii. We report design guidelines for the implementation of the classification algorithm, the size of the training set, and the training procedure.