Behavioral Insights on Influence of Manual Action on Object Size Perception

Annalisa Bosco, Patrizia Fattori
Department of Pharmacy and Biotechnology
University of Bologna
Bologna, Italy
patrizia.fattori@unibo.it

Abstract— **Visual perception is one of the most advanced function of human brain. The study of different aspects of human perception currently contributes to machine vision applications. Humans estimate the size of objects to grasp them by perceptual mechanisms. However, the motor system is also able to influence the perception system. Here, we found modifications of object size perception after a reaching and a grasping action in different contextual information. This mechanism can be described by the Bayesian model where action provides the likelihood and this latter is integrated with the expected size (prior) derived from the stored object experience (Forward Dynamic Model). Beyond the action-modulation effect, the knowledge of subsequent action type modulates the perceptual responses shaping them according to relevant information required to recognize and interact with objects. Cognitive architectures can be improved on the basis of these processings in order to amplify relevant features of objects and allow to robot/agent an easy interaction with them.**

Keywords—visual perception, object recognition, motor output, human functions, context information.

I. Introduction

The majority of machine vision and object recognition systems today apply mechanistic or deterministic template matching, edge detection or color scanning approach for identifying different objects in the space and also to guide embodied artificial intelligent systems to interaction with them. However, fine disturbances in the workspace of a robot can lead to failures, and thus slow down their performance in identification, recognition, learning and adapting to noisy environment, compared to human brain. To go beyond these limitations robots with intelligent behavior must be provided with a processing architecture that allows them to learn and reason about responses to complex goals in a complex world.

The starting point for the development of such intelligent systems is the study of human behavior. Humans frequently estimate the size of objects to grasp them. In fact, when performing an action, our perception is focused towards object visual properties that enable us to execute the action successfully. However, the motor system is also able to influence perception, but only few studies reported evidence for action-induced visual perception modifications related to hand movements [1–4]. For example, the orientation perception is enhanced during preparation of grasping action compared with a pointing for which object orientation is not important [5,6]. This "enhanced perception" is triggered by the intention to grasp and is important to examine objects with the maximum possible accuracy. If we consider the effects of action execution on visual perception of object features, there is ample evidence for visual perception changes in the oculomotor system, but little is known about the perceptual changes induced by different types of hand movements. In order to evaluate the influence of different hand movement on visual perception, we tested a feature-specific modulation on object size perception after a reaching and a grasping action in different contexts.

II. Materials and Methods

A total of 16 right-handed subjects (11 females and 5 males, ages 21–40 years; with normal or corrected-to-normal vision) took part in the experiment. The experiment was performed by two groups of participants. One group of 8 subjects performed the Prior knowledge of action type experiment (PK condition) and the other group (8 participants) performed the No prior knowledge of action type (NPK condition). All subjects were naive to the experimental purpose of the study and gave informed consent to participate in the experiment. Procedures were approved by the Bioethical Committee of the University of Bologna and were in accordance with the Declaration of Helsinki.

A. Apparatus and Setup

Participants were seated in an environment with dim background lighting and viewed a touchscreen monitor (ELO IntelliTouch, 1939L), which displayed target stimuli within a visible display of 37.5 X 30.0 cm. To stabilize head position, the participants placed their heads on a chin rest located 43 cm from the screen, which resulted in a visual field of 50 x 40 deg. The display had a resolution of 1152 X 864 pixels and a frame rate of 60 Hz (15,500 touch points/cm^2). For stimulus presentation, we used MATLAB (The MathWorks) with the Psychophysics toolbox extension [7]. The stimuli were white, red and green dots with a radius of 1.5 mm and 10 differently sized white, red and green bars all 9 mm large and whose length was: 30, 33.6, 37.2, 40.8, 44.4, 48, 51.6, 55.2, 58.8, 62.4 mm. Hand position was measured by a motion capture system (VICON, 460; frequency of acquisition 100 Hz), which follows the trajectory of the hand in three dimensions by recording infrared light reflection on passive markers.

Participants performed 10 blocks of 10 trials each. Each trial consisted of three successive phases: *Pre-size perception*, *Reaching or Grasping movement*, *Post-size perception* (Fig. 1). In Pre-size perception and Post-size perception phases (phases 1 and 3), a white or green central fixation target stayed on the screen for 1 s; then, a white or green bar was presented, for 1 s, 12 deg on the left or on the right side of the central fixation target and, after an acoustic signal, it disappeared. The participants were required to manually indicate the perceived horizontal size of the bar. All participants indicated the bar sizes by keeping the hand within the starting hand position square and the distance between subject eyes. In the Reaching or Grasping movement phase (phase 2), after 1 s, the white or green central fixation point was followed by a bar identical for position and size to that of phases 1 and 3. Participants were required to perform a reaching (closed fist) or grasping action (extension of thumb and index fingers to "grasp" the extremities of the bar) towards the bar after the acoustic signal, respectively. The type of actions was instructed by the colors of the stimuli (fixation point and bar). In fact, if the color of the stimuli was white, participants were required to perform a reaching movement whereas, if the color was green, they were required to perform a grasping movement. In PK condition, the color of fixation points and bars was white or green in all three phases of trial and in this way the participants knew in advance (from phase 1) which action type was required in the movement phase (phase 2). In the NPK condition, the sequence of the three phases was identically structured as in the PK condition, but we changed colors of fixation points and bars from white/green to red in phases 1 and 3. The color of stimuli during phase 2 remained white or green according to the movement type, reaching or grasping respectively. By this color manipulation, participants could not know in advance the successive action type.

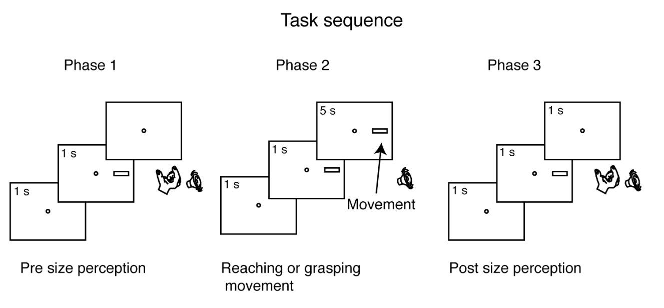

Fig. 1. *Task sequence. Circle = fixation point, Rectangle = stimulus, Hand = size indication by manual report, Speaker = acoustic signal to respond.*

B. *Data analysis*

After data collection, finger position data were interpolated at 1000 Hz, then data were run though a fifth-order Butterworth low-pass filter [8]. For data processing and analysis, we wrote custom software in MATLAB to compute the distance between index and thumb markers during the pre- and post-manual estimation phases. Grip aperture was calculated considering trial intervals in which the velocities of the index and thumb markers remained <5 mm/s [8]. Grip aperture was defined as maximum distance within this interval. To evaluate the effect of different hand movement on size perception, we compared the manual perceptual responses before the movements with those after the movements by using two-tailed t-test with independent samples.

To evaluate the magnitude of the effect of NPK and PK conditions on perceptual responses before the movement we calculated the average difference between the two responses and we compared the responses between the two conditions by a t-test analysis. We extracted relevant features from the perceptual responses before the movement and we used them to predict the NPK and PK conditions. For this purpose, we performed a linear-discriminant analysis (LDA-based classifier), as implemented in Statistics and Machine Learning toolbox (Matlab). Pre movement manual responses of NPK and PK conditions were vertically concatenated to build the feature space composed by 958 trials. Fivefold cross-validation was performed by using the 80% of trials for training and the 20% for testing the data, so to ensure that the classifier was trained and tested on different data. Specifically, the classifier was trained on the training subset and the obtained optimal decision criteria was implemented on the testing subset. The prediction results were obtained for this testing subset. This procedure was repeated 5 times, so that all trials were tested and classified basing on models learned from the other trials. The prediction results for all the trials were taken together to give an averaged prediction result with standard deviation. We considered statistically significant the accuracies which standard deviations did not cross the theoretical chance level of 50%. We used a LDA classifier as decoder of the two conditions. LDA finds linear combination of features that characterizes or separates two or more classes of objects or event [9,10]. In fact, LDA explicitly attempts to model the difference between the classes of data. For all statistical analyses the significant criterion was set to $P < 0.05$.

III. RESULTS

We assessed the effects of action execution on perceptual responses comparing the single subject responses before the movement with those after the movement and calculating the difference between these. Fig. 2 shows these differences in grey color for reaching movement on the horizontal axis compared with those of grasping movement on vertical axis. Filled and empty circles are referred to PK and NPK condition, respectively. The majority of subjects fell below the diagonal suggesting that they corrected the perceptual estimation after the grasping movements with respect to the reaching movement. In particular, they perceived significantly smaller the bars after a grasping movement with respect to a reaching movement ($P < 0.05$). The averaged differences in PK and NPK conditions are reported in Fig. 2 as black and white dots, respectively. Both dots are below the diagonal suggesting that, globally, subjects perceived smaller after a grasping action compared with a reaching action.

To analyze the effect of the NPK and PK conditions on size perception, we focused the analyses on manual size reports before the movement execution (Pre size perception phase). We computed the difference between the Pre size perception reports in PK condition and the Pre size perception reports in NPK condition. This difference allowed to highlight the amount of change in size perception in the two conditions tested. As it is shown in Fig. 3A, we found that the amount of change in reaching was -11.89 mm ±0.98 mm and in grasping -11.36 mm ±1.08 mm, and in both cases, they were significantly deviated from baseline (t-test, $P < 0.05$). Generally, the subjects tended to perceive smaller the sizes presented in the condition where they were aware about the subsequent action (PK condition) compared with the condition where they were uncertain about the successive movement (NPK condition). To evaluate whether the strength of this effect was due to a perceptual bias or to different neural processings, we used a LDA decoder to classify the manual responses according to the NPK and PK condition (see Material and Methods). In other words, we checked whether we were able to predict the PK and NPK conditions from perceptual responses before the movement execution, as this technique represents a powerful method to reconstruct experimental conditions and functional movements from neural responses using different types of classifiers [11,12]. Fig. 3A shows decoding results as confusion matrix and the corresponding mean accuracy expressed in percentage. We found a good correlation between the real conditions and the decoded conditions, as it is illustrated in Fig. 3B. The accuracies of decoding were significantly higher of 50% (66,8% for PK and 60.54% for NPK) as shown in Fig. 3C.

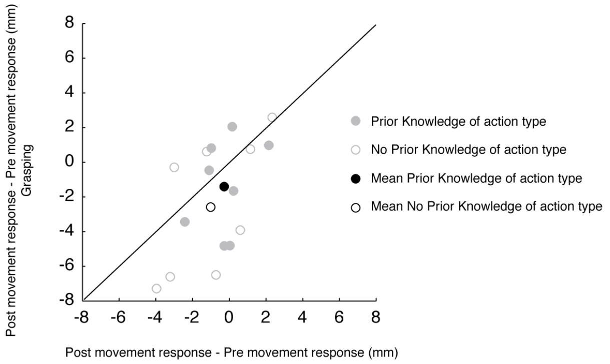

Fig. 2. *Differences between perceptual responses before and after the movement. Filled grey dots are differences in PK condition and empty grey dots are differences in NPK condition. Black and white dots are the mean differences in PK and NPK conditions, respectively.*

IV. Discussion

In the present study, we found direct evidence for a perceptual modification of a relevant feature as object size before and after the execution of two types of hand movement. These changes depended on two factors: the knowledge of the subsequent action type and the type of action executed. Changes in perception were sharpened after a grasping action compared with a reaching. Specifically, subjects perceived objects smaller after a grasping movement than after a reaching movement. The study of action effects exerted by the skeletomotor system on perception has been focused on the evidence that relevant features of objects, such as size or orientation, prime the perceptual system in order to execute a more accurate subsequent grasping movement.

Indeed, Gutteling et al. [5] demonstrated an increased perceptual sensitivity to object orientation during a grasping preparation phase. The effect of action-modulated perception has also been shown to facilitate visual search for orientation. Bekkering and Neggers [2] analysed the performance of subjects that were required to grasp or point to an object of a certain orientation and color among other objects. They demonstrated that fewer saccadic eye movements were made to wrong orientations when subjects had to grasp the object than point to it. Recently, Bayesian theory has been applied to formalize processes of cue and sensorimotor integration [13,14]. According to this view, the nervous system combines prior knowledge about object properties gained through former experience (prior) with current sensory cues (likelihood), to generate appropriate object properties estimations for action and perception. Hirsinger and coworkers [15], by application of a size-weight illusion paradigm, found that the combination of prior and likelihood for size perception were integrated in a Bayesian way. Their model consisted in a Forward Dynamic Model (FDM) that represented the stored object experience. The FDM output was the experience-based expected size and was referred as the prior. The prior then was integrated with the likelihood, which represented the afferent sensory information about object size. A feedback loop with a specified gain provides the FDM with the final estimate of size, which serves as learning signal for adapting object experience. In the present study, we can apply a similar model for size perception after an action execution. In our case, the objects were visual, not real objects and no haptic feedback was given after the execution of movement. So, the likelihood was represented by the matching of the fingers with the outer border of objects with/or the proprioceptive signals coming from the hand posture that are integrated with the prior.

We found that the knowledge of action type was a factor modulating size perception. In fact, subjects perceived smaller the bars during the condition where they knew the subsequent action (PK) compared with the other condition where they did not know the subsequent action (NPK) for both reaching and grasping. A further demonstration of that was related to the possibility to predict with significant accuracy (>50%) the two conditions from perceptual responses before movement (see Fig. 3B-C). This approach is typical for neural responses and represents a novelty for this type of behavioral variables. The significance of these results is in line with evidence from behavioral research suggesting that motor planning processes increase the weight of visual inputs.

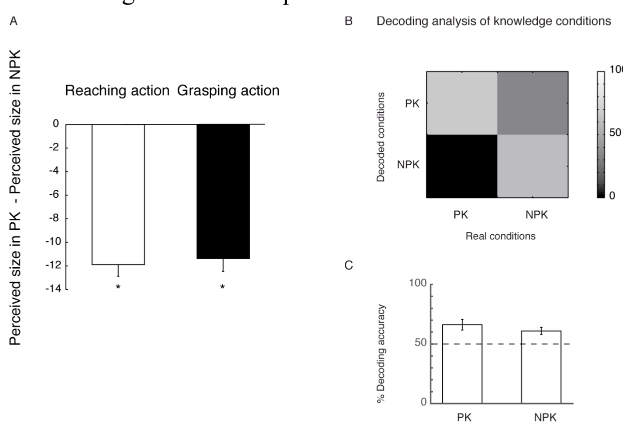

Fig.3. *A, Mean differences of perceptual responses between PK and NPK conditions in reaching and grasping. B, Confusion matrix of decoding results. C, Mean decoding accuracy for classification of NPK and PK conditions. Error bars are standard deviation. *P<0.05, significant level.*

Hand visual feedback has been found to have a greater impact on movement accuracy when subjects prepare their movements with the prior knowledge that vision will be available during their reaches [16,17]. More interestingly, motor preparation facilitates the processings of visual information related to the target of movement. Similarly to Gutteling et al. [5] for object orientation, Wykowska et al. [18] reported that the detection of target size was facilitated during the planning of grasping but not during the planning of pointing. All these studies show the capacity of the brain to modulate the weight of visual inputs and provide an illustration of the importance of the context in visual information processing. In line with all these studies, our findings suggest that the knowledge or not of subsequent movement type defines a context that modulates the perceptual system. When subjects knew the subsequent movement, the perceptual system was within a definite context and perceived object smaller, scaling the measures according to hand motor abilities. In the other case, subjects were in an uncertain context about the successive action, and the perceptual system used different rules to scale the size reports. In both cases, the defined and undefined context can be predicted. All the mechanisms described in the present study could implement models of cognitive architecture of vision-based reaching and grasping of objects located in the peripersonal space of a robot/agent. Additionally, the evidence that the perceptual system is dynamically modulated by contextual information about subsequent movement type can be used to improve cognitive architectures. For example one or multiple focus of attention signals can be sent to the object representation of robot/agent in order to amplify relevant features and at the same time inhibits distractors.

ACKNOWLEDGMENT

We thank F. Daniele for helping in the data collection and in data analysis. This work was supported by Firb 2013 N. RBFR132BKP (MIUR) and by the Fondazione del Monte di Bologna e Ravenna.

REFERENCES

[1] Craighero, L., Fadiga, L., Rizzolatti, G., Umiltà C. Action for Perception A Motor - Visual Attentional Effect. J Exp Psychol Hum Percept Perform. 1999;25: 1673–.

[2] Bekkering H, Neggers SFW. Visual search is modulated by action intentions. Psychol Sci a J Am Psychol Soc / APS. 2002;13: 370–374. doi:10.1111/j.0956-7976.2002.00466.x

[3] Hannus A, Cornelissen FW, Lindemann O, Bekkering H. Selection-for-action in visual search. Acta Psychol (Amst). 2005;118: 171–191. doi:10.1016/j.actpsy.2004.10.010

[4] Fagioli S, Hommel B, Schubotz RI. Intentional control of attention: Action planning primes action-related stimulus dimensions. Psychol Res. 2007;71: 22–29. doi:10.1007/s00426-005-0033-3

[5] Gutteling TP, Kenemans JL, Neggers SFW. Grasping preparation enhances orientation change detection. PLoS One. 2011;6. doi:10.1371/journal.pone.0017675

[6] Gutteling TP, Park SY, Kenemans JL, Neggers SFW. TMS of the anterior intraparietal area selectively modulates orientation change detection during action preparation. J Neurophysiol. 2013;110: 33–41. doi:10.1152/jn.00622.2012

[7] Brainard DH. The Psychophysics Toolbox. Spat Vis. 1997;10: 433–436. doi:10.1163/156856897X00357

[8] Bosco A, Lappe M, Fattori P. Adaptation of Saccades and Perceived Size after Trans-Saccadic Changes of Object Size. J Neurosci. 2015;35: 14448–14456. doi:10.1523/JNEUROSCI.0129-15.2015

[9] Fisher RA. The use of multiple measurements in taxonomic problems. Ann Eugen. 1936;7: 179–188.

[10] McLachlan GJ. Discriminant analysis and statistical pattern recognition. Wiley Intersci. 2004;. ISBN 0-4.

[11] Schaffelhofer S, Agudelo-Toro A, Scherberger H. Decoding a wide range of hand configurations from macaque motor, premotor, and parietal cortices. J Neurosci. 2015;35: 1068–81. doi:10.1523/JNEUROSCI.3594-14.2015

[12] Townsend BR, Subasi E, Scherberger H. Grasp movement decoding from premotor and parietal cortex. J Neurosci. 2011;31: 14386–98. doi:10.1523/JNEUROSCI.2451-11.2011

[13] Körding KP, Wolpert DM. Bayesian decision theory in sensorimotor control. Trends in Cognitive Sciences. 2006. pp. 319–326. doi:10.1016/j.tics.2006.05.003

[14] Van Beers RJ, Wolpert DM, Haggard P. When feeling is more important than seeing in sensorimotor adaptation. Curr Biol. 2002;12: 834–837. doi:10.1016/S0960-9822(02)00836-9

[15] Hirsiger S, Pickett K, Konczak J. The integration of size and weight cues for perception and action: Evidence for a weight-size illusion. Exp Brain Res. 2012;223: 137–147. doi:10.1007/s00221-012-3247-9

[16] Zelaznik HZ, Hawkins B, Kisselburgh L. Rapid visual feedback processing in single-aiming movements. J Mot Behav. 1983;15: 217–236. doi:10.1080/00222895.1983.10735298

[17] Elliott D, Allard F. The utilization of visual feedback information during rapid pointing movements. Q J Exp Psychol A Hum Exp Psychol. 1985;37: 407–425. doi:10.1080/14640748508400942

[18] Wykowska A, Schubö A, Hommel B. How you move is what you see: action planning biases selection in visual search. J Exp Psychol Hum Percept Perform. 2009;35: 1755–1769. doi:10.1037/a0016798