NeuroImage 125 (2016) 668–680

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Encodings of implied motion for animate and inanimate object categories in the two visual pathways

Zhengang Lu [a,1], Xueting Li [a,b,1], Ming Meng [a,*]

[a] Dartmouth College, Hanover, NH, USA
[b] Renmin University of China, Beijing, China

ARTICLE INFO

Article history:
Received 17 June 2015
Accepted 20 October 2015
Available online 24 October 2015

Keywords:
Implied motion
motion speed
animate-inanimate
visual pathways
univariate averaged BOLD
MVPA

ABSTRACT

Previous research has proposed two separate pathways for visual processing: the dorsal pathway for "where" information vs. the ventral pathway for "what" information. Interestingly, the middle temporal cortex (MT) in the dorsal pathway is involved in representing implied motion from still pictures, suggesting an interaction between motion and object related processing. However, the relationship between how the brain encodes implied motion and how the brain encodes object/scene categories is unclear. To address this question, fMRI was used to measure activity along the two pathways corresponding to different animate and inanimate categories of still pictures with different levels of implied motion speed. In the visual areas of both pathways, activity induced by pictures of humans and animals was hardly modulated by the implied motion speed. By contrast, activity in these areas correlated with the implied motion speed for pictures of inanimate objects and scenes. The interaction between implied motion speed and stimuli category was significant, suggesting different encoding mechanisms of implied motion for animate-inanimate distinction. Further multivariate pattern analysis of activity in the dorsal pathway revealed significant effects of stimulus category that are comparable to the ventral pathway. Moreover, still pictures of inanimate objects/scenes with higher implied motion speed evoked activation patterns that were difficult to differentiate from those evoked by pictures of humans and animals, indicating a functional role of implied motion in the representation of object categories. These results provide novel evidence to support integrated encoding of motion and object categories, suggesting a rethink of the relationship between the two visual pathways.

Introduction

To survive in a dynamic world, the sensitivity of the human visual system for detecting motion cues is a critical evolutionary advantage. This motion sensitivity is so delicate that motion perception can occur even when no physical motion is presented but only implied (Freyd, 1983). Converging neurophysiological and neuroimaging evidence indicates that the middle temporal (MT) area plays a central role in the perception of motion (Maunsell and Van Essen, 1983; Van Essen et al., 1981; Tootell et al., 1995a; Dupont et al., 1994; Orban et al., 1995; Born and Bradley, 2005). Interestingly, recent studies have further shown that the MT responds to not only physical motion (e.g. moving dots and gratings), but also dynamic information contained in still photographs when motion is not presented, which is known as implied motion (Zeki et al., 1993; David and Senior, 2000; Kourtzi and Kanwisher, 2000a, 2000b; Senior et al., 2000; Peuskens et al., 2005; Krekelberg et al., 2005; Kim and Blake, 2007). Apparently, the dynamic information in implied motion does not come from direct motion signals, but instead may be implied from the formation of global patterns (Krekelberg et al., 2003; Krekelberg et al., 2005) or from object categorization and the knowledge about how animate and inanimate objects move (Kourtzi and Kanwisher, 2000a).

However, the visual system is widely known to consist of two separate pathways: the dorsal "where" pathway encodes spatial location and motion-related information, whereas the ventral "what" pathway encodes shape and form information (Mishkin et al., 1983). Representing implied motion must rely on the integration of object and motion information, as the encoding of implied motion is predicting "where" things will be based on "what" stimuli are. Thus, how the two visual pathways encode implied motion remains an important question to be addressed. On the one hand, if the MT in the dorsal pathway is not involved in object categorization/representation, the known MT activity corresponding to implied motion might reflect feedback processes occurring elsewhere in the brain (Kourtzi and Kanwisher, 2000a). Converging neuroimaging results have demonstrated that animate-inanimate distinction is a major dimension for how the visual system represents objects in the temporal cortex (Chao et al., 1999; Downing et al., 2006; Wiggett et al., 2009; Martin, 2007; Grill-Spector and Weiner, 2014; Kriegeskorte et al., 2008; Cichy et al., 2014; Sha et al., 2014). It is possible

---

* Corresponding author at: Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755.
E-mail address: ming.meng@dartmouth.edu (M. Meng).

[1] These authors contributed equally.

http://dx.doi.org/10.1016/j.neuroimage.2015.10.059
1053-8119/

that such encodings of animacy in the ventral pathway feedback to the dorsal pathway, and thus inform the MT to process how animate and inanimate objects move. Nonetheless, how the encoding of animacy may interact with the encoding of implied motion is unknown. For example, it is not clear whether the ventral pathway may encode a flying plane differentially than a plane on the ground, as no one has directly examined whether implied motion is encoded in the ventral pathway. Correspondingly, it is unknown whether in the representational space of object categories in the ventral pathway, the similarity between a flying plane and a bird may be greater than the similarity between a motionless plane (inanimate object) and a bird (animate object).

On the other hand, it has been reported that the MT is also involved in the processing of shape properties of moving objects in addition to motion (Kourtzi et al., 2001). Such findings suggested that the MT may play a functional role more than just computing the direction and speed of motion (Born and Bradley, 2005). However, it is unclear whether the MT may also encode object categories. For physically moving stimuli, MT neurons have been shown to be selective for motion speed as well as motion direction (Lagae et al., 1993; Maunsell and Van Essen, 1983; Perrone and Thiele, 2001). It remains possible that the encoding of motion (and implied motion) in the MT may in turn mediate encodings of object categories, such as differentiating animate vs. inanimate objects. In summary, the relationship between how the brain encodes motion and how the brain encodes object/scene categories remains unclear.

To address this question, the present fMRI study examines how the object and motion information contained in implied motion stimuli may be encoded at multiple areas in the two visual pathways. A 4 × 5 factorial design is used, as stimuli include four animate/inanimate categories (humans, animals, objects, and scenes) that induce implied motion at five levels of perceived motion speed (Fig. 1). Note that stimulus category is a discrete nominal variable, whereas motion speed is a continuous ratio variable. Convergent evidence suggests that, although averaged BOLD responses may correlate with continuous ratio variables such as contrast, decoding of the representations corresponding to categorical variables is better accomplished through multivoxel analyses of activation patterns (Haxby et al., 2001; Cox and Savoy, 2003; Guo and Meng, 2015). In the present study, we localized regions of interests (ROIs) in both dorsal and ventral visual pathways. We then analyzed both averaged BOLD activity and patterns of activity in these ROIs corresponding to when observers were shown the stimuli consisting of 4 categories and 5 levels of implied motion speed. Since speed tuning in MT neurons is well known (Lagae et al., 1993; Maunsell and Van Essen, 1983; Perrone and Thiele, 2001), we hypothesize that if the brain encodes implied motion in a similar way as encoding physical motion, fMRI activity in at least the MT may be found to correlate with implied dynamic levels of motion speed. Moreover, whether or not implied motion speed relies on stimulus category to elicit fMRI activity is a separate open question. The processing of biological motion involves different mechanisms from the encoding of non-biological motion (Beauchamp et al., 2002; Beauchamp et al., 2003; Grossman et al., 2000; Peuskens et al., 2005). It is possible that the encoding of implied motion speed in biological motion (i.e. implied motion from pictures of humans and animals) may be different from the encoding of implied motion speed in non-biological motion. In addition to ANOVAs, linear trend analysis is conducted to examine whether increased implied motion speed for each stimulus category may lead to increased BOLD activity in different ROIs. Finally, through multivariate pattern analyses, should implied motion be encoded through categorical animate-inanimate distinctions, we would expect to decode stimulus category based on the activity patterns of ROIs not only in the ventral pathway but also in the dorsal pathway including the MT. However, should the encoding of motion be independent of stimulus category, fMRI activity corresponding to implied motion would be found to correlate only with the ordinal implied dynamic levels of motion speed regardless of the categorical animate-inanimate distinction.

Material and Method

Participants

Thirteen students (5 females) at Dartmouth College volunteered to participate in the experiment. They were compensated for their time. All participants had normal or corrected-to-normal visual acuity. None of the participants had been informed the purpose of the experiment and none had participated in any other experiments using the same set of stimuli. Data from two participants were excluded from further analysis because of excessive head motion. This research was approved by the Committee for the Protection of Human Subjects at Dartmouth College. All participants gave written informed consent.

Materials

A total of 112 pictures, containing four categories (humans, animals, objects, and scenes), were collected from the Internet. They were equated in mean luminance and contrast by using MATLAB and the SHINE toolbox (Willenbockel et al., 2010). A separate group of 24 participants first were shown all of the 112 pictures one by one, so they were made aware of the content and dynamic range of the pictures. Then, they were asked to judge whether or not each picture presented was dynamic, and finally were asked to evaluate the dynamic level of implied motion speed of each picture on a Likert scale considering the whole range of all the pictures. Based on the ratings, pictures varying in five levels of implied motion speed (level 0 = static to level 4 = most dynamic) were selected for each of the four categories, resulting in 20 pictures in total. These pictures were then scaled to a fixed size of 500 by 400 pixels (12.5° × 10 ° of visual degree) for the following fMRI experiment.

MRI acquisition

MRI experiments were conducted with a 3.0 T Philips Intera Achieva scanner (Philips Medical Systems, Bothell, WA) using a 32 channels SENSE (SENSEitivity Encoding) head coil at Dartmouth Brain Imaging Center. A high-resolution T1-weighted 3D-MPRAGE anatomical scan was acquired for each participant (FOV = 240 mm, TR = 8.2 ms, TE = 3.8 ms, flip angel = 8°, voxel size = 1 mm × 1 mm × 1 mm, reconstruction matrix = 256 × 256, 222 slices). To measure BOLD contrast, 35 slices/volume parallel to the anterior commissure/posterior commissure line were acquired using standard gradient-echo echo-planar imaging (EPI) sequence (FOV = 240 mm, TR = 2000 ms, TE = 35 ms, flip angle = 90°, slice thickness = 3 mm, slice gap = 0.3 mm, inplane resolution = 3 × 3 mm). All functional images were acquired in an interleaved slice order. Stimuli were presented to participants via a Panasonic DT-4000U DLP projector. The rear-projection screen was positioned at the rear of the scanner and viewed with a mirror mounted to the head coil. The width and height of the projected screen were 45.7 cm and 34.3 cm (1024 × 768 pixels) respectively. The distance between the mirror and projected screen was 97.8 cm. The distance between participants' eyes and mirror was approximately 12.7 cm.

Procedures

All subjects participated in two sessions of fMRI experiments on two separate days (at least one week apart). For each session, EPIs were collected in 6 experimental runs with a slow event-related design and 2 regions of interests (ROI) localizer runs with a block design. In addition, a block-designed MT localizer run was conducted during the second session. During each of the 6 experimental runs, a fixation cross (about 1°) was always presented at the center of the screen, and 146 volumes (sets of axial images) were collected in 292 s. Each run began and ended with 12 s of a fixation-only resting period and contained 20 trials of stimulus presentation in a random sequence. For each trial, a mirror version of a stimulus picture was presented side by side