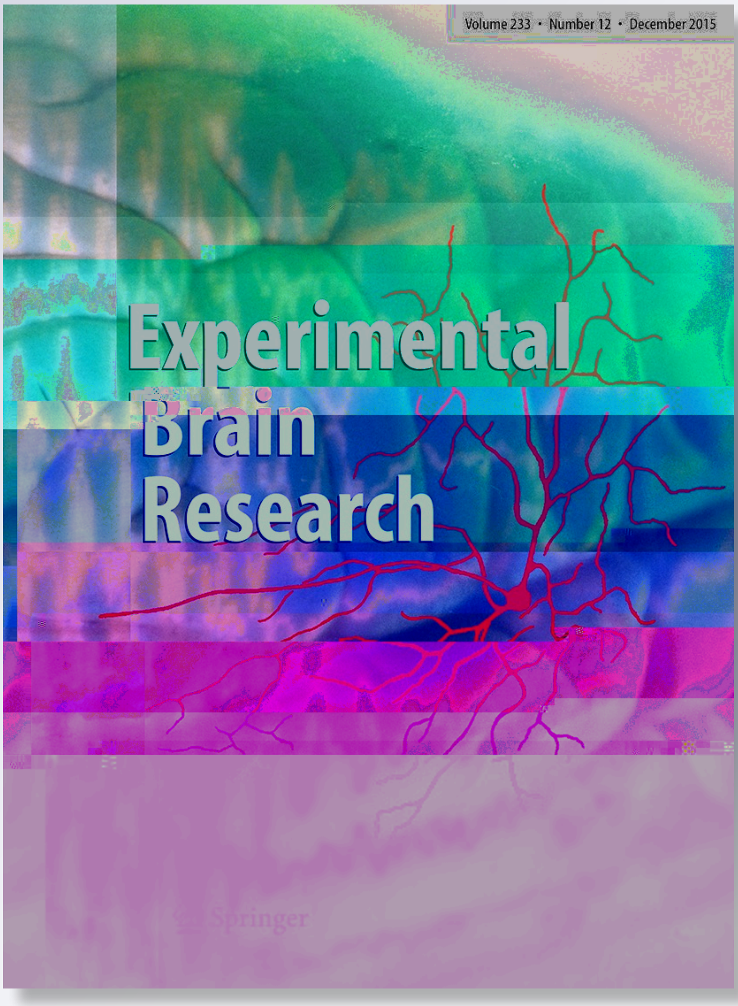
Position shifts of fMRI-based population receptive fields in human visual cortex induced by Ponzo illusion

Dongjun He, Ce Mo, Yizhou Wang & Fang Fang

Experimental Brain Research

ISSN 0014-4819
Volume 233
Number 12

Exp Brain Res (2015) 233:3535-3541
DOI 10.1007/s00221-015-4425-3



Exp Brain Res (2015) 233:3535–3541
DOI 10.1007/s00221-015-4425-3

RESEARCH ARTICLE

Position shifts of fMRI-based population receptive fields in human visual cortex induced by Ponzo illusion

Dongjun He[1,2,3,4] · **Ce Mo**[3] · **Yizhou Wang**[5] · **Fang Fang**[1,2,3,4]

Received: 5 June 2015 / Accepted: 18 August 2015 / Published online: 28 August 2015

Abstract Ponzo illusion is a well-known perceptual phenomenon in which the perceived sizes of visual objects are altered by visual depth cues created by converging lines at the horizon. One possible neural mechanism of the Ponzo illusion is the receptive field position shifts of V1 neurons, as supported by a recent monkey electrophysiological study (Ni et al. in Curr Biol 24(14):1653–1658, 2014). Here, we used fMRI-based population receptive field (pRF) mapping technique in combination of psychophysics to investigate this idea. We found that, relative to the close apparent depth in a 3D scene, the far apparent depth in the scene caused the pRF positions of voxels in V1–V3 to shift toward the fovea, in line with subjects' percept of the Ponzo illusion. Moreover, the pRF position shift in V1 significantly correlated with the magnitude of the Ponzo illusion across individual subjects. Our findings thus provide evidence for the close association between the perceived object size and the pRF position shift in human visual areas, especially in V1, lending further support for the receptive field position shift explanation for the Ponzo illusion.

Keywords Ponzo illusion · Population receptive field · Functional brain imaging · Psychophysics

Introduction

Size illusion is a perceptual phenomenon of the misjudgment of physical sizes of visual objects induced by contextual cues. One of the most representative types of size illusion is the Ponzo illusion in which two identical objects appear to be of different sizes when placed at different visual depths created by converging lines at the horizon (Leibowitz et al. 1969). Previous studies (Fang et al. 2008; Murray et al. 2006) have identified a close relationship between the perceived object size during the Ponzo illusion and the spatial distribution pattern of neural activity in primary visual cortex (V1). The researchers presented two rings at close and far apparent depths in a 3D scene. When subjects fixated its center, the far ring appeared to be larger and occupy a more eccentric portion of the visual field, relative to the close ring. Using functional magnetic resonance imaging (fMRI), they found that, relative to the close ring, the spatial distribution of V1 activity induced by the far ring also shifted toward the peripheral representation of the visual field, demonstrating that the Ponzo illusion was reflected in the spatial distribution of neural activity in V1. Furthermore, it was found that the magnitude of the Ponzo illusion was dependent on the attention focus of the observer. The cortical representation of the ring size was evidently modulated by the visual depth cues only when attention was directed to the rings. Together, these

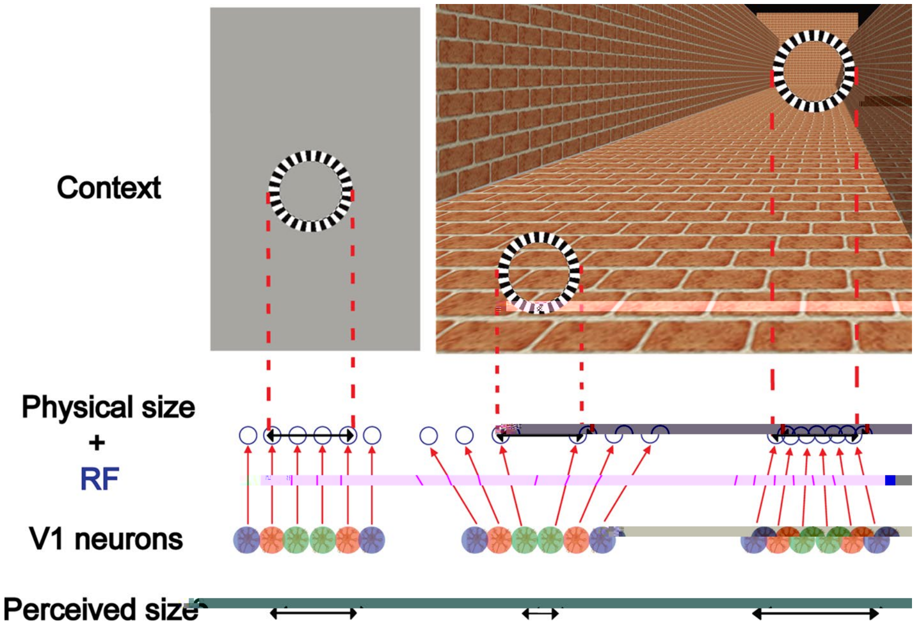


Fig. 1 A model of perceived size representation in early visual cortex. The perceived size is determined by the spatial distribution of activated neurons. The ring stimulus presented with no context (*left*) activates neurons through whose RFs it passes. The ring that appears to be close is perceived as small, consistent with outward RF shifts and a small cortical representation (*middle*). Conversely, the ring that appears to be far is perceived as large, consistent with inward RF shifts and a large cortical representation (*right*)

findings suggest that the Ponzo illusion is mediated by the topographic dynamics of V1 neural activity modulated by feedback signal from higher cortical areas.

However, it is still unclear what mechanisms drive the shifts of the spatial distribution of V1 activity during the Ponzo illusion. Neurons in V1 are characterized by their precise retinotopic organization. Their receptive field positions could serve to encode the spatial properties (e.g., location and size) of visual inputs. To represent the perceptually larger ring, neurons with more peripheral receptive fields need to be recruited to encode the perceived size of the ring. Hence, a possible neuronal mechanism in V1 for generating the Ponzo illusion is the position shift of V1 neuronal receptive fields initiated by feedback signal because visual depth information in a 3D scene is highly complex and extracting the information is beyond the local processing capacity of V1 neurons with limited receptive field sizes. In this case, the far ring elevates the firing rate of peripheral neurons as a result of an inward shift of the receptive fields of the neurons toward the fovea, leading to more peripheral V1 activity (Fig. 1). However, it is traditionally believed that the receptive field positions of V1 neurons are mainly determined by the pattern of feed forward connections from the retina to the cortex independent of current task demands and attention states (Alonso et al. 2001; Lund et al. 2003; Reid and Alonso 1995). Although a recent study reported neuronal receptive field position shifts in macaque primary visual cortex during the Ponzo illusion (Ni et al. 2014), so far there is no such evidence in human subjects because of the rare opportunity of intracranial recording in human early visual cortex (Mukamel and Fried 2012).

One approach to address this question is the non-invasive fMRI-based population receptive field (pRF) mapping technique (Dumoulin and Wandell 2008). This technique is based on the assumption that the joint receptive field of the neuronal population within a single voxel can be characterized by a two-dimensional Gaussian function with three parameters, x_0, y_0, and σ, where x_0 and y_0 determine the position (i.e., center) of the joint receptive field in the visual field and σ (i.e., dispersion of the Gaussian) determines the size of the receptive field. By fitting the predicted signal based on this model to the BOLD signal time course, the pRF position and size parameters can be estimated for individual voxels, thus providing a characterization of the receptive field properties of neuronal populations across the visual cortex. This method has been shown to reconstruct the cortical visual field map more accurately than conventional phase-encoded retinotopic mapping methods and produce pRF size estimates that agree well with electrophysiological measurements in monkey and human subjects (Kok and de Lange 2014).

In the current study, we used the fMRI-based pRF mapping technique in combination of psychophysics to investigate whether the Ponzo illusion is mediated by systematic shifts of pRF positions in human early visual cortex. The objective of the psychophysical experiment was to measure the magnitude of the Ponzo illusion, and the pRF mapping experiment was designed to measure the pRF properties in human early visual areas during the Ponzo illusion. We used similar checkered ring stimuli and 3D visual scene as those in Fang et al. (2008) (Fig. 2a). We hypothesized that, relative to the close ring, the depth cues in the 3D scene would cause systematic inward shifts of the pRF positions

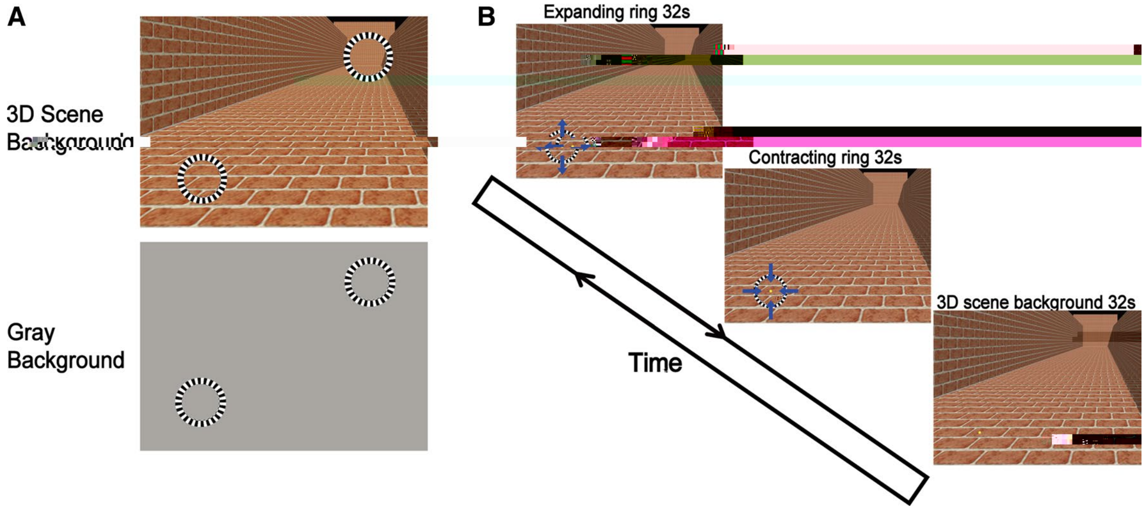


Fig. 2 Stimuli and experimental procedure. **a** Stimuli. One stimulus was a rendered 3D scene of a hallway and walls with two physically identical rings: one ring was at a close apparent depth and the other was at a far apparent depth. In the other stimulus, the two rings were presented against a gray background. **b** fMRI experimental procedure. Expanding and contracting rings were presented at either the close or far apparent depth for estimating pRF parameters in V1–V3

of the voxels responding to the far ring area, leading to a larger perceived size of the far ring.

Materials and methods

Subjects

A total of nine human subjects (6 male, 18–28 years old) were paid to take part in the study. All of them participated in both the psychophysical and fMRI experiments. All subjects were naïve to the purpose of the study. They were right-handed, reported normal or corrected-to-normal vision, and had no known neurological or visual disorders. They gave written informed consent, and our procedures were approved by the Human Subject Review Committee at Peking University.

Psychophysical experiment

The psychophysical experiment was performed in a separate session outside the magnet. Visual stimuli were presented on an IIYAMA monitor (model HM204DT, size 22 inches) with a spatial resolution of 1024 × 768 and a refresh rate of 85 Hz. Subjects viewed the stimuli from a distance of 73 cm with their head stabilized on a chin rest. We used a method of constant stimuli to measure the magnitude of the Ponzo illusion. We presented two rings in a rendered three-dimensional (3D) scene of a hallway and walls: one ring was at a close ("front") apparent depth and the other one was at a far ("back") apparent depth (Fig. 2a). In a trial, subjects were given 2 s of free viewing time before the rings disappeared and then were required to indicate which ring looked larger. The size of the front ring was fixed. Its inner and outer radii were 1.96° and 2.52°, respectively. The size of the back ring varied from trial to trial. Its inner and outer radii were randomly selected from five pairs of values: 1.40°/1.96°, 1.68°/2.24°, 1.96°/2.52°, 2.24°/2.80°, and 2.52°/3.08°. Note that all the front and back rings had a fixed width (0.56°). Subjects completed 200 trials in total, with 40 trials for each back ring size. In addition to the 3D scene condition, we also presented the rings on a gray background with no contextual cue and measured the perceived size difference between the rings, serving as a baseline for comparison (Fig. 2a). For each of the two background conditions, we plotted the percentage of trials in which subjects indicated that the back ring was perceived to be larger than the front ring as a function of the physical size of the back ring. To quantitatively measure the illusion size, we fit the psychometric values at the five physical sizes with a cumulative normal function. We interpolated the data to find the point of subjective equality (PSE) at which the front and back rings appeared to have the same size.

fMRI experiments

Retinotopic visual areas (V1, V2, and V3) were defined by a standard phase-encoded method developed by Sereno et al. (1995) and Engel et al. (1997), in which subjects viewed a rotating wedge and an expanding ring that created traveling waves of neural activity in visual cortex.

An independent block-design run was performed to identify the voxels or ROIs in the retinotopic areas responding to the front and back rings when subjects fixated at the center of the rings. The run contained eight stimulus blocks of 12 s, interleaved with eight blank blocks of 12 s. The stimulus was a full-contrast flickering checkered ring (inner radius 1.42°, outer radius 3.10°).

To estimate pRF parameters, we measured hemodynamic response function (HRF) for each subject in a separate run. This run contained 12 trials. In each trial, a full-contrast flickering checkered disc with a radius of 10.94° was presented for 2 s, followed by a 30-s blank interval. The HRF was measured by fitting the convolution of a 6-parameter double-gamma function with a 2-s boxcar function to the BOLD response elicited by the disc. In both the localizer run and the HRF estimation run, subjects performed a color discrimination task at fixation point to maintain fixation and control attention.

The pRF mapping experiment consisted of 8 functional runs of 192 s (Fig. 2b). In half of the runs, a series of flickering checkered rings with 16 different sizes (inner/outer radius of the smallest ring 0°/0.56°, inner/outer radius of the largest ring 4.2°/4.76°, step size 0.28°) were presented at the front ring position, while in the other half, the ring stimuli were presented at the back ring position. There were two cycles in each run. Each cycle included a 32-s expanding block in which the rings were presented in a small-to-large order, a 32-s contracting block in which the rings were presented in a large-to-small order, and a 32-s background block in which only the 3D scene (without the rings) was presented. Subjects were required to attend to the ring stimuli and detect occasional, brief pauses of flickering.

MRI data acquisition

All MRI experiments were performed on a 3T Siemens Prisma MRI scanner at the Center for MRI Research at Peking University. MRI data were acquired with a 20-channel phase-array head coil. In the scanner, the stimuli were back-projected via a video projector (refresh rate 60 Hz, spatial resolution 1024 × 768) onto a translucent screen placed inside the scanner bore. Subjects viewed the stimuli through a mirror mounted on the head coil. The viewing distance was 73 cm. BOLD signals were measured using an echo-planner imaging (EPI) sequence (TE 30 ms, TR 2000 ms, flip angle 90°, acquisition matrix size 76 × 76, FOV 152 × 152 mm^2, slice thickness 2 mm, gap 0 mm, number of slices 33, slice orientation axial). The fMRI slices covered the occipital lobe and a small part of the cerebellum. Before the functional scans, a T1-weighted high-resolution 3D structural data set was acquired for each participant in the same session using a 3D-MPRAGE sequence (voxel size 1 × 1 × 1 mm^3). Subjects underwent two sessions, one for retinotopic mapping and ROI localization, and the other for estimating pRF parameters.

MRI data analysis

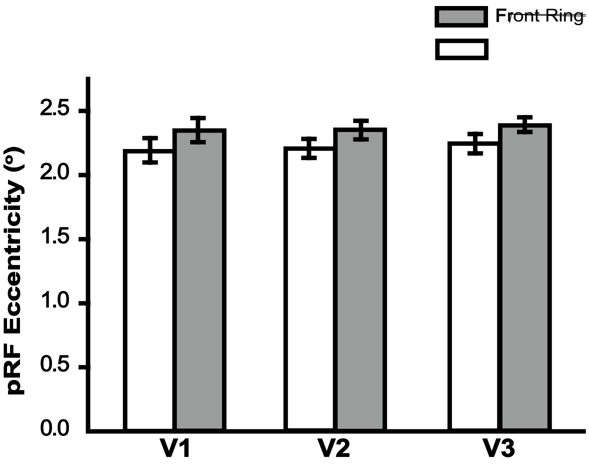
MRI data were processed using BrainVoyager QX (Brain Innovations, Maastricht, The Netherlands) and custom scripts in Matlab (Mathworks Inc.). The anatomical volume in the retinotopic mapping session was transformed into the Talairach space and then inflated using BrainVoyager QX. Functional volumes in all sessions were preprocessed, including 3D motion correction, linear trend removal, and high-pass filtering (cutoff frequency 0.015 Hz) using BrainVoyager QX. No subject exhibited excessive head movement (<2 mm in translation, <0.5° in rotation) within any fMRI session. The functional volumes were then aligned to the anatomical volume in the retinotopic mapping session and transformed into the Talairach space. The first 6 s of BOLD signals were discarded to minimize transient magnetic saturation effects.

A general linear model (GLM) procedure was used for ROI localization. The ROIs in V1–V3 were defined as cortical areas or voxels that responded more strongly to the flickering checkered ring than the blank screen ($p < 10^{-5}$, uncorrected) in the localizer run. We estimated the pRF of each voxel in the ROIs using the method proposed by Dumoulin and Wandell (2008). In this analysis, a predicted BOLD signal is calculated from the known visual stimulus parameters and a model of the underlying neuronal population. The model of the neuronal population consisted of a two-dimensional Gaussian pRF, with parameters x_0, y_0, and σ, where x_0 and y_0 are the coordinates of the center of the receptive field, and σ indicates its spread (standard deviation), or size. All parameters were stimulus-referred, and their units were degrees of visual angle. These parameters were adjusted to obtain the best possible fit of the predicted to the actual BOLD signal. Statistical analyses were performed with voxels whose pRF model could explain at least 10 % of the variance of the raw data.

Results

Psychophysical results

We used a method of constant stimuli to measure the magnitude of the Ponzo illusion. The front ring was presented with a fixed size. Its inner and outer radii were 1.96° and 2.52°, respectively. The back ring was presented with various sizes. We searched the PSE of the back ring at which the back and front rings appeared to have the same size. We found that, when the back ring had the inner and outer radii



of 1.81° and 2.37° (5.95 % of the outer radius of the front ring), the two rings were perceived to be equally large. The magnitude of the Ponzo illusion was quantified as the size difference between the two rings when the PSE was identified (mean ± sem 0.15° ± 0.02°). The illusion size was robust and significant [t(8) = 8.68, $p < 0.001$] (Fig. 3a). It is noteworthy that, when the front and back rings were presented on a gray background, the illusion completely vanished (0.07 % of the outer radius of the front ring), suggesting that the Ponzo illusion observed in the 3D scene condition was not simply due to the locations of the rings (Fig. 3a).

fMRI results

We estimated the pRF positions of the voxels corresponding to the front and back rings, respectively. The estimated positions were converted to eccentricities and then averaged across voxels in V1, V2, or V3 (Fig. 3b). In V1, the mean pRF eccentricities associated with the front and back rings were 2.34° and 2.18°, respectively. This finding demonstrated that the far depth information, relative to the near depth information, caused an inward pRF shift toward the fovea [t(8) = 4.09, $p < 0.005$]. A similar pattern was found in the other two visual areas. In V2, the mean pRF eccentricities associated with the front and back rings were 2.35° and 2.20°, respectively [t(8) = 4.33, $p < 0.005$]. In V3, the mean pRF eccentricities associated with the front and back rings were 2.38° and 2.24°, respectively [t(8) = 5.15, $p < 0.001$].

To further examine the behavioral relevance of the observed pRF position shifts, we calculated the Pearson's correlation coefficient between the magnitude of the Ponzo illusion and the pRF position shifts in V1–V3 (Fig. 4). We found a significant positive correlation in V1 ($r = 0.92$, $p < 0.001$), but not in V2 ($r = 0.569$, $p = 0.11$, 32 % variance explained) and V3 ($r = 0.64$, $p = 0.063$, 41 % variance explained). The correlation coefficient in V1 was (marginally) significantly larger than those in V2 ($p < 0.05$) and V3 ($p = 0.07$), suggesting that V1 was more closely associated with the Ponzo illusion than V2 and V3.

Discussion

Our study found that, relative to the close apparent depth in the 3D scene, the far apparent depth caused the pRF positions of voxels in V1–V3 to shift toward the fovea, consistent with the perceived size of the front and back rings. Moreover, the pRF position shift in V1 significantly correlated with the magnitude of the Ponzo illusion across individual subjects. It should be noted that, the observed effects are unlikely due to difference in spatial attention or eye movement, because subjects performed the same task with the front and back rings.

Fang et al. (2008) revealed that the Ponzo illusion could alter the spatial pattern of V1 neural activity and the perceived larger ring activated more peripheral areas in V1, compared with the perceived smaller ring. To represent the perceptually larger ring, neurons with more peripheral receptive fields need to be recruited to encode the perceived size of the back ring. The observed inward shift of the pRFs toward the fovea entails more peripheral neurons to represent the perceived size of the back ring, leading to more peripheral V1 activity. Thus, our finding here is in line with Fang et al. (2008). Using extracellular recordings, Ni et al. (2014) also found a similar receptive field shift of V1 neurons. Our study here examined not only V1, but also extrastriate areas V2 and V3, thus providing a more comprehensive characterization of neural changes during the Ponzo illusion. Notably, our observation of the systematic receptive field shifts in V1–V3 is in line with the neurological zoom system account for perceptual size constancy proposed in an early study (Marg and Adams 1970). Using the technique of implanted micro-electrode recording, the researchers found that the neuronal receptive fields in human early visual cortex were "magnified" in response to an incoming object such that the perceived object size was not altered by the increased retinal image. Together, these studies make headways toward unraveling the neural mechanisms of the Ponzo illusion.

Consistent with our results, recent studies also demonstrated that the topography of V1 neural activity closely mirrored the perceived object size that was significantly affected by spatiotemporal contexts (e.g., physical distance or the size of visual adaptor), even when the retinal input remains unchanged (Pooresmaeili et al. 2013; Sperandio et al. 2012). In the study conducted by Sperandio et al. (2012), they presented subjects with a constant light stimulus and manipulated the perceived size of the stimulus afterimage by varying the viewing distance of a back screen holding the afterimage. They found that the spatial distribution of V1 activity expanded to the more eccentric representation of the visual field as the perceived size of the afterimage increased with the viewing distance, suggesting neuronal populations in V1 encode the size information of the perceived rather than the physical visual input. In another study performed by Pooresmaeili et al. (2013), using a visual adaptation paradigm, they found that the spatial extent of V1 activation closely resembled the perceived size of the testing stimulus that was altered by the size of adapting stimulus. Moreover, other recent studies also demonstrated that neural activity in V1 could encode other perceived stimulus properties (e.g., brightness and shape; Boyaci et al. 2007; Michel et al. 2013). Together, these findings provide strong evidence of the critical role of V1 in encoding perceived (rather than physical) visual information, which is usually believed to take place in high visual areas.

Across individual subjects, the pRF shift in V1 exhibited a stronger coupling with the magnitude of the Ponzo illusion than those in V2 and V3. One possible explanation is that the integration of size and distance information for size–distance scaling is dependent on top-down modulation that is better captured by the BOLD responses in striate than in extrastriate areas (Sperandio and Chouinard 2015). This speculation is in accordance with a recent anatomical finding (Schwarzkopf et al. 2011). Schwarzkopf et al. (2011) found that the magnitude of the Ponzo illusion was negatively correlated with the surface area of V1 but not with that of V2 and V3. Together, these findings reveal the functional and anatomical bases underlying the interindividual difference in susceptibility to the Ponzo illusion. How does our functional finding relate to Schwarzkopf et al.'s anatomical finding? Horizontal connections in V1 provide an effective means for individual neurons to integrate and assess visual information over a large portion of the visual field beyond their own receptive fields (Gilbert and Wiesel 1989), which might be critical for the receptive field shift of V1 neurons observed here. As the area of V1 surface increases, the horizontal connections become less effective as the information has to travel longer distances for interneuron communication, leading to smaller receptive field shifts. Hence, we speculate that horizontal connections in V1 might serve as an important predictor of individual susceptibility to the Ponzo illusion.

In sum, we show that the neural coding of perceived object sizes during the Ponzo illusion is implemented in the human brain via systematic shifts of pRF positions throughout early visual areas, which provides the non-invasive model-based fMRI evidence in human subjects echoing previous invasive electrophysiological findings in monkey subjects.

Acknowledgments This work was supported by the Ministry of Science and Technology of China (2015CB351800) and the National Natural Science Foundation of China (31230029 and 31421003). We thank the staff at the Center for MRI Research at Peking University for technical support.

References

Alonso JM, Usrey WM, Reid RC (2001) Rules of connectivity between geniculate cells and simple cells in cat primary visual cortex. J Neurosci 21(11):4002–4015

Boyaci H, Fang F, Murray SO, Kersten D (2007) Responses to lightness variations in early human visual cortex. Curr Biol 17(11):989–993

Dumoulin SO, Wandell BA (2008) Population receptive field estimates in human visual cortex. Neuroimage 39(2):647–660

Engel SA, Glover GH, Wandell BA (1997) Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cereb Cortex 7(2):181–192

Fang F, Boyaci H, Kersten D, Murray SO (2008) Attention-dependent representation of a size illusion in human V1. Curr Biol 18(21):1707–1712

Gilbert CD, Wiesel TN (1989) Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. J Neurosci 9(7):2432–2442

Kok P, de Lange FP (2014) Shape perception simultaneously up-and downregulates neural activity in the primary visual cortex. Curr Biol 24(13):1531–1535

Leibowitz H, Brislin R, Perlmutrer L, Hennessy R (1969) Ponzo perspective illusion as a manifestation of space perception. Science 166(3909):1174–1176

Lund JS, Angelucci A, Bressloff PC (2003) Anatomical substrates for functional columns in macaque monkey primary visual cortex. Cereb Cortex 13(1):15–24

Marg E, Adams JE (1970) Evidence for a neurological zoom system in vision from angular changes in some receptive fields of single neurons with changes in fixation distance in the human visual cortex. Experientia 26(3):270–271

Michel MM, Chen Y, Geisler WS, Seidemann E (2013) An illusion predicted by V1 population activity implicates cortical topography in shape perception. Nat Neurosci 16(10):1477–1483

Mukamel R, Fried I (2012) Human intracranial recordings and cognitive neuroscience. Ann Rev Psychol 63:511–537

Murray SO, Boyaci H, Kersten D (2006) The representation of perceived angular size in human primary visual cortex. Nat Neurosci 9(3):429–434

Ni AM, Murray SO, Horwitz GD (2014) Object-centered shifts of receptive field positions in monkey primary visual cortex. Curr Biol 24(14):1653–1658

Pooresmaeili A, Arrighi R, Biagi L, Morrone MC (2013) Blood oxygen level-dependent activation of the primary visual cortex predicts size adaptation illusion. J Neurosci 33(40):15999–16008

Reid RC, Alonso JM (1995) Specificity of monosynaptic connections from thalamus to visual cortex. Nature 378(6554):281–283

Schwarzkopf DS, Song C, Rees G (2011) The surface area of human V1 predicts the subjective experience of object size. Nat Neurosci 14(1):28–30

Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ et al (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268(5212):889–893

Sperandio I, Chouinard PA (2015) The mechanisms of size constancy. Multisens Res 28(3–4):253–283

Sperandio I, Chouinard PA, Goodale MA (2012) Retinotopic activity in V1 reflects the perceived and not the retinal size of an afterimage. Nat Neurosci 15(4):540–542