

RESEARCH ARTICLE

# A Neural Model of Distance-Dependent Percept of Object Size Constancy

Jiehui Qian<sup>1</sup>, Arash Yazdanbakhsh<sup>2\*</sup>

**1** Department of Psychology, Sun Yat-Sen University, Guangzhou, China, **2** Department of Psychological & Brain Sciences, Center for Computational Neuroscience and Neural Technology, Boston University, Boston, Massachusetts, United States of America

\* [yazdan@bu.edu](mailto:yazdan@bu.edu)



**OPEN ACCESS**

**Citation:** Qian J, Yazdanbakhsh A (2015) A Neural Model of Distance-Dependent Percept of Object Size Constancy. PLoS ONE 10(7): e0129377. doi:10.1371/journal.pone.0129377

**Editor:** Suliann Ben Hamed, Centre de Neurosciences Cognitive, FRANCE

**Received:** December 3, 2014

**Accepted:** May 5, 2015

**Published:** July 1, 2015

**Copyright:** © 2015 Qian, Yazdanbakhsh. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper.

**Funding:** This work has been supported in part by the Center of Excellence for Learning in Education, Science and Technology (CELEST, NSF SBE-0354378), Office of Naval Research (ONR N00014-11-1-0535), Air Force Office of Scientific Research (AFOSR, FA9550-12-1-0436), and the Fundamental Research Funds for the Central Universities (14wkpy20). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Size constancy is one of the well-known visual phenomena that demonstrates perceptual stability to account for the effect of viewing distance on retinal image size. Although theories involving distance scaling to achieve size constancy have flourished based on psychophysical studies, its underlying neural mechanisms remain unknown. Single cell recordings show that distance-dependent size tuned cells are common along the ventral stream, originating from V1, V2, and V4 leading to IT. In addition, recent research employing fMRI demonstrates that an object's perceived size, associated with its perceived egocentric distance, modulates its retinotopic representation in V1. These results suggest that V1 contributes to size constancy, and its activity is possibly regulated by feedback of distance information from other brain areas. Here, we propose a neural model based on these findings. First, we construct an egocentric distance map in LIP by integrating horizontal disparity and vergence through gain-modulated MT neurons. Second, LIP neurons send modulatory feedback of distance information to size tuned cells in V1, resulting in a spread of V1 cortical activity. This process provides V1 with distance-dependent size representations. The model supports that size constancy is preserved by scaling retinal image size to compensate for changes in perceived distance, and suggests a possible neural circuit capable of implementing this process.

## Introduction

Humans make stable judgments about an object's actual size despite changes in its retinal size with distance, which is known as size constancy phenomenon. Previous research proposed a size-distance invariance hypothesis (SDIH) to explain the size constancy. It states that some function of retinal size combines multiplicatively with perceived distance to obtain the perceived size of an object [1–4]. Many size illusions, such as the Ponzo and moon illusions, are suggested to be based on this size–distance relationship [5–9]. The SDIH has long been proposed, however, its underlying neural mechanisms remain unclear.

Single cell recordings in awake and anesthetized monkeys confirm the existence of distance-dependent size tuned cells along the ventral pathway from visual cortical area V1, V2 and V4

**Competing Interests:** The authors have declared that no competing interests exist.

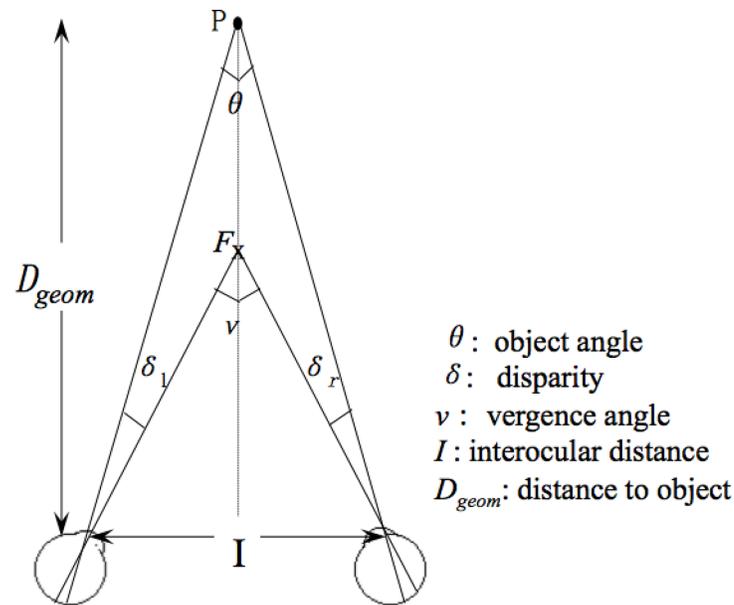
[10] leading to inferotemporal (IT) cortex [11]. In particular, Dobbins et al. [10] found that a large number of cells in V1, V2, and V4 preferred the same retinal image size, but varied their firing rates with different viewing distances. Among these cells, some showed a monotonic increase in mean firing rate with decreasing distance (nearness cells), some with increasing distance (farness cells), and a few that are distance-independent. These results suggest that size tuned cells whose activities are scaled by distance are common in the visual cortex. In addition, recent functional magnetic resonance imaging (fMRI) studies [12–14] demonstrate that an object’s perceived size modulates its cortical retinotopic representation. A distant object would appear to be larger than a closer object with the same retinal image size. Surprisingly, the apparently larger object causes activation at a more eccentric locus in the primary visual cortex, compared to the apparently smaller object. In other words, the same visual angle projected on the retina could occupy different proportions of V1, depending on the perceived size of the corresponding object. Although these findings contradict the traditional view that retinotopic mapping in V1 is precise and hardwired, emerging evidence supports that visual processing in V1 depends on both retinal image and distance information [15]. The distance may be signaled by feedback of three-dimensional (3D) spatial representation from other brain regions, such as the lateral intraparietal cortex (LIP) [16].

Distance estimation is crucial to size constancy. Egocentric distance, i.e., the distance of an object from an observer, can be estimated by horizontal disparity (HD) combined with viewing distance. Viewing distance usually refers to the distance to the fixation point. It can be recovered by depth cues, such as vergence angle and vertical disparities [17, 18]. Besides disparity and vergence, there are many other depth cues, such as motion parallax, occlusion, familiar size, and linear perspective, which could contribute to distance perception. However, given the limited viewing conditions used in the single cell and imaging studies, it is assumed that disparity and vergence play a dominant role in distance perception. In this article, we only focus on these two depth cues. If an object is directly in front of the observer, its distance,  $D_{geom}$ , can be recovered by a function of HD,  $\delta$ , vergence,  $v$ , and interocular distance,  $I$  (see S1 Table for the full list of variables). According to the geometrical layout in Fig 1, the equation is as follows:

$$D_{geom} = \frac{I}{2 \tan\left(\frac{v-\delta}{2}\right)} \tag{1}$$

where  $I$  normally range from 6 mm to 7 mm. When disparity increases, Eq (1) initially underestimates and then overestimates the perceived distance [19–20]. By modifying  $\delta$ , Eq (1) can approximate perceived distance quite accurately [21]. This modified  $\delta$  is called corrected disparity, see S1 File for details.

Estimation of perceived distance depends on disparity, while disparity estimation may also depend on distance. Many studies have shown that disparity tuned cells are modulated by viewing distance. Trotter et al. [22, 23] found that in alert, behaving monkeys, the responses of a large majority of disparity-tuned neurons in V1 were distance-dependent. They suggested that extraretinal signals, probably vergence or accommodation, could be integrated with disparity early in the visual processing pathways for 3D spatial representation. Other studies supported that vergence can be used as a reliable cue for distance perception [24–26]. Since it is known that cells in V1 are sensitive to HD [27–30], disparity signals in our proposed neural model arise from V1. While studies have suggested multiple loci in the brain responsible for vergence control [31–33], we select the frontal eye fields (FEF) to provide vergence signal in our model [34, 35]. These two depth cues give rise to distance perception. We predict that area MT may be responsible for integrating vergence and disparity signals. Although so far we lack



**Fig 1. Geometry for stereopsis.** "F" marks the fixation point;"P" marks the position of an object.  $D_{geom}$  indicates the distance from object, and  $I$  indicates interocular distance. The horizontal disparity,  $\delta$ , is equal to the difference of the vergence angle,  $v$ , and the object angle,  $\theta$ .  $\delta = \delta_r + \delta_l$ , where  $\delta_r$  and  $\delta_l$  correspond to the disparity on the left and right retina respectively.

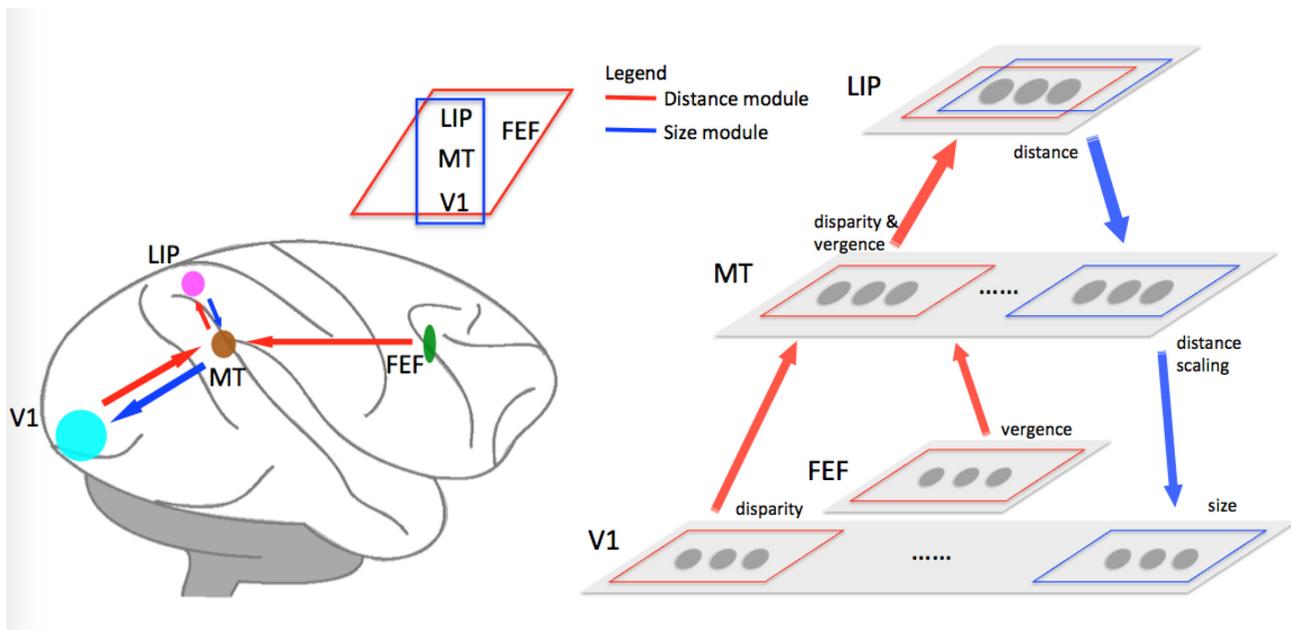
doi:10.1371/journal.pone.0129377.g001

direct neurophysiological evidence to confirm this point, monosynaptic connections have long been found from V1 to MT [36, 37] and FEF to MT [35].

Single cell recordings found that LIP provides a distributed representation of egocentric space [16, 38–40]. It seems that LIP receives inputs from MT and MST, generating a 3D spatial representation that acts as a premotor signal for directing saccades [16]. Based on these finding, we hypothesizes that MT feeds the integrated depth information forward to LIP, where a distance map is constructed. Then, LIP feeds the distance information back to MT [39, 41], and even further back to V1, where it regulates responses of size tuned cells to achieve a distance-dependent size representation.

We propose a neural model that simulates the results from Dobbins et al. [10] and Sperandio et al. [14]. The model comprises of a distance module and a size module (Fig 2). In the distance module, first a horizontal disparity signal arises from V1 and vergence signal arises from FEF; then MT neurons gate disparity tuning by vergence; finally, the information in MT neurons pass on to LIP to construct a 3D spatial representation. In the size module, the distance information obtained from LIP provides a feedback signal to gain modulate neural responses of size tuned cells and the spread of cortical activity in V1. Table 1 summarizes the brain areas involved in our model and their proposed functions. However, these proposed functions need to be tested.

Although there are neural models proposed to simulate the cortical representation of egocentric distance [21], no previous modeling work has clearly demonstrated the neural mechanisms of size constancy. Here, our model simulates the neural responses of distance-dependent size tuned cells in V1, based on the current findings of the neurophysiology and fMRI studies. Our model supports that size constancy could be preserved by scaling retinal image size to compensate for changes in viewing distance, and suggests a neural circuit of how this process is achieved.



**Fig 2. Proposed brain areas and their connections for distance-dependent size perception.** Red arrows indicate pathways involved in distance perception: fibers carrying signals from V1 (disparity) and FEF (vergence) terminate in MT; MT then projects to LIP, where a map of distance information is constructed. Blue arrows indicate pathways involved in size perception: the distance information from LIP provides feedback signals to MT, which further projects to V1, modulating neural responses of size tuned cells and cortical activity spread in V1. Thus, a distance-dependent size representation is achieved.

doi:10.1371/journal.pone.0129377.g002

## Methods

### Model overview

Our model consists of two sequential stages: first, the distance module where a distance map is created; second, the size module where size tuning is regulated by distance. In the first stage (Fig 2, red arrow), cells in area V1 code for HD using Gaussian functions, while cells in area FEF code for vergence using sigmoidal functions. MT cells integrate the outputs from both V1 and FEF by means of a set of basis functions; the outputs of MT cells feed forward to cells in LIP to construct a distance map. In the second stage (Fig 2, blue arrow), the distance information from LIP feeds back to MT for distance scaling and then back to V1, modulating the activity of size tuned cells. A size representation based on distance therefore can be constructed.

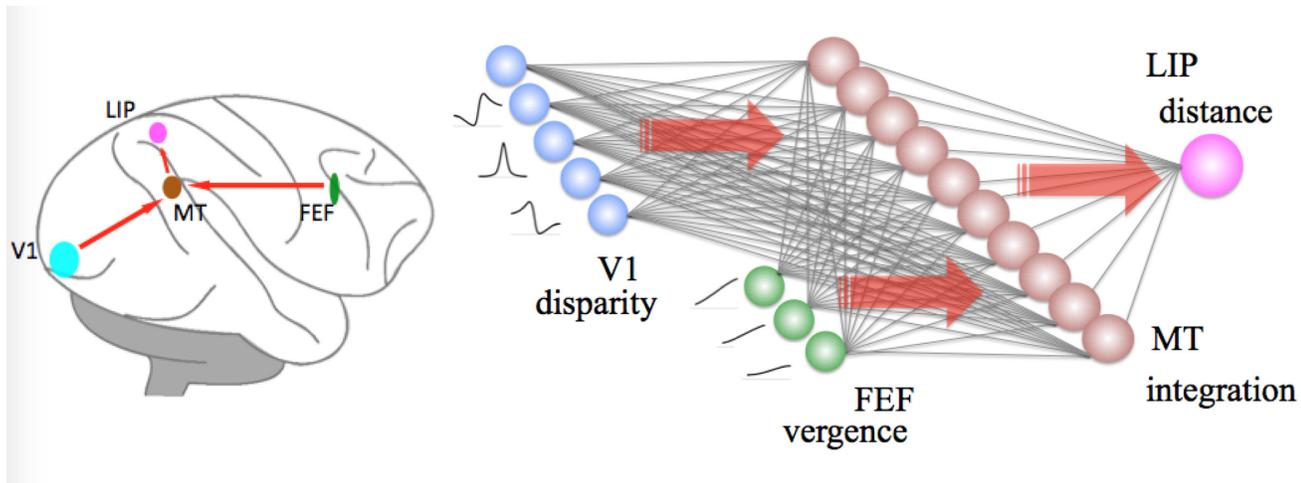
### Distance Module

We propose that the distance module consists of bidirectional connections from V1 and FEF to MT and further to LIP (Fig 3). Area MT can be deemed as a layer of processing units

**Table 1. Brain areas and their proposed functions in the model.**

Model areas	Distance module (Ref.)	Size module (Ref.)
V1	disparity coding [22]	size coding [10, 14]
FEF	vergence coding [34]	—
MT	disparity & vergence (prediction)	distance scaling (prediction)
LIP	distance coding [32]	distance

doi:10.1371/journal.pone.0129377.t001



**Fig 3. Distance module.** Disparity selective cells (see variable  $a$  in Eqs (4)–(6)) in V1 are gain-modulated and simulated by Gaussian functions. Vergence selective cells (see variable  $v$  in Eq (7)) in FEF are simulated by sigmoidal functions. At the level of MT (see variable  $B$  in Eq (3)), vergence and disparity are integrated. At the level of LIP, a distance map (see variable  $D$  in Eq (2)) is estimated from a linear combination of MT outputs.

doi:10.1371/journal.pone.0129377.g003

integrating inputs of disparity signals from V1 and vergence signals from FEF. It serves as an intermediate transformation between the inputs from V1 and FEF, and outputs to LIP.

The final output of the distance module is a distance map,  $D(\delta, v)$ , in LIP. It is estimated from a linear combination of the outputs of MT:

$$D = \sum_{i=1}^n w_i B_i(\delta, v) \quad (2)$$

where  $w_i$  is the connection weight from each cell in MT to LIP.  $B_i(\delta, v)$  corresponds to the responses of gain-modulated MT neuron,  $i$ , which is specified as follows:

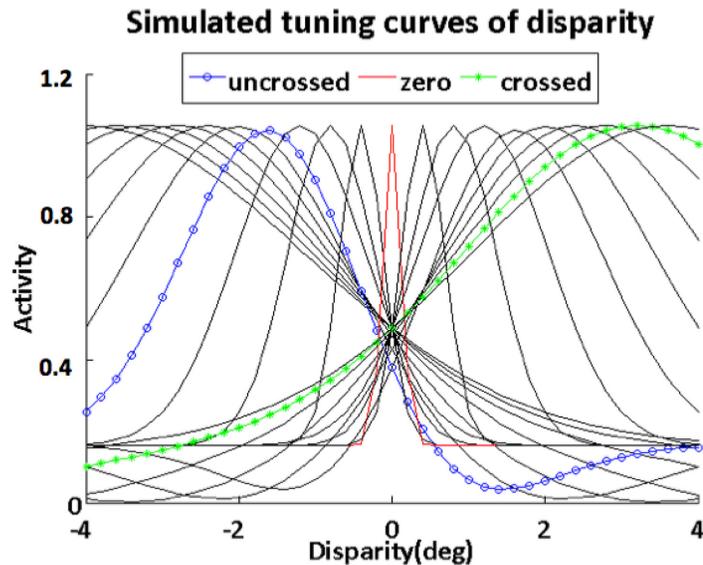
$$B_i(\delta, v) = a_i(\delta) z_i(v) \quad (3)$$

where  $a_i(\delta)$  and  $z_i(v)$  indicate the tuning curves of disparity- and vergence- selective cells, respectively (see S1 Table for the full list of variables). It is known that disparity-selective cells in V1 are gain-modulated, and can be categorized into one of the six cell types [28, 30] (see Fig 9 of [30]): tuned excitatory (TE)/tuned inhibitory (TI) cell, which gives maximal/minimal responses at zero disparity; tuned near (TN)/tuned far (TF) cell, which has similar disparity tuning function as TE cells but peaks at negative/positive disparity; far (FA)/near (NE) cell, which activates over a wide range of positive/negative disparities. The disparity tuning of V1 neurons are generally described using either Gabor [42–44] or Gaussian functions [21, 45]. In our model, the tuning profile of V1 cells are defined by Gaussians for mathematical simplicity: Near/tuned near (with preferred disparity,  $\delta_i$ , within  $[-4^\circ, -1.5^\circ]$ ) cells,

$$a_i(\delta) = A_1 e^{-\frac{(\delta-\delta_i)^2}{\sigma_i^2}} - A_2 e^{-\frac{(\delta-(\delta+\sigma_i^2))^2}{\sigma_i^2}} + A_3 \quad (4)$$

Far/tuned far (with preferred disparity,  $\delta_i$ , within  $[1.5^\circ, 4^\circ]$ ) cells,

$$a_i(\delta) = A_1 e^{-\frac{(\delta-\delta_i)^2}{\sigma_i^2}} - A_2 e^{-\frac{(\delta-(\delta-\sigma_i^2))^2}{\sigma_i^2}} + A_3 \quad (5)$$



**Fig 4. Simulated tuning curves of disparity-selective cells.** The cell tuning curves in V1 [28, 30] are simulated by subtracting Gaussians. Each curve corresponds to the response of a single cell and exhibits a similar response pattern as Fig 9 in Gonzalez and Perez [30]. For example, the red dashed line resembles the TE cell's response, which gives maximal responses at zero disparity; the blue open-circle line resembles the TN cell's response, which has similar disparity tuning function as TE cells but peaks at negative disparity; the green star line resembles the FA cell's response, which activates over a wide range of positive disparities. TF and NE cells are simulated but not marked in the figure. TI cells are not included in the simulation, since their responses could be simulated by assigning negative output weights to the TE cells.

doi:10.1371/journal.pone.0129377.g004

Tuned excitatory units (with preferred disparity,  $\delta_i$ , within  $[-1.5^\circ, 1.5^\circ]$ ) cells,

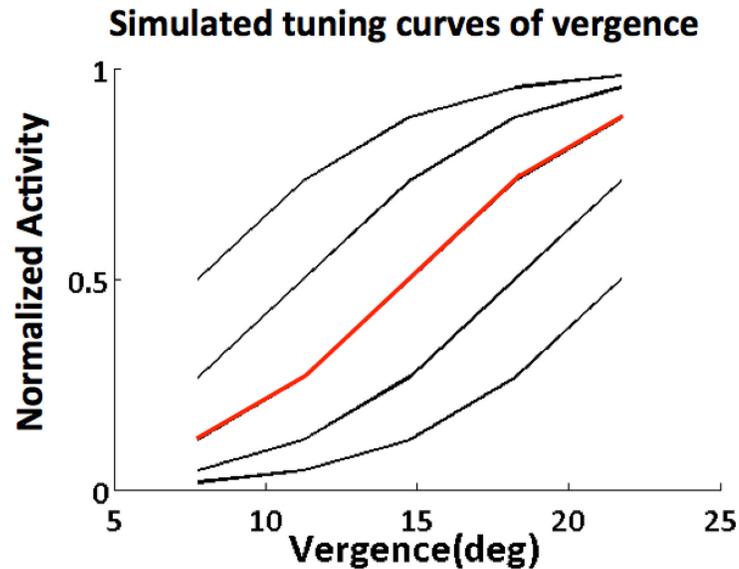
$$a_i(\delta) = A_1 e^{-\frac{(\delta - \delta_i)^2}{\sigma_i^2}} \tag{6}$$

where  $\delta_i$  is the peak of response for a given curve, and  $\sigma_i$  describes the width of the tuning curve for the full list of variables), which is chosen to equal the absolute value of the disparity corresponding to the peak response, except for the curves whose peaks are within the disparity range  $[-10, 10']$ , for which  $\sigma_i$  is set to 10 arcmin. Forty disparity tuned cells were assigned. Their tuning curves,  $a_i(\delta)$ , are shown in Fig 4 (every other cell responses are shown). Resembling Fig 9 in Gonzalez and Perez [30], our model tuning curves exhibit similar response patterns of tuned near, tuned far, tuned excitatory, near and far cells found in primates. Parameters,  $A_1$ ,  $A_2$ , and  $A_3$  were chosen such that each cell types' tuning curve  $a_i$  would have approximately the same maximum value.

Vergence tuning curves in FEF,  $z_i(v)$ , are modeled as sigmoids:

$$z_i(v) = \frac{1}{1 + e^{-\frac{v - v_i}{T_i}}} \tag{7}$$

where  $v$  is the vergence angle,  $v_i$  and  $T_i$  are the thresholds and the slopes of the sigmoids, respectively. The threshold controls the position of the sigmoid and the vergence axis, and the slope controls the steepness of the sigmoid. We chose five different thresholds, therefore obtaining five vergence tuning curves. The vergence tuning curves are shown in Fig 5. Resembling cell responses recorded from the primate visual cortex (refer to Fig 3b in Gamlin [34]), our model tuning curves show that cell activity increases as vergence angle increases.



**Fig 5. Simulated tuning curves of vergence cells.** The modeled tuning curves of multiple vergence cells in FEF are simulated by sigmoids. Each curve corresponds to the response of a single cell to vergence. Similar to Fig 3b in Gamlin [34], our model tuning curves show that cell activity increases as vergence angle increases.

doi:10.1371/journal.pone.0129377.g005

Disparity and vergence signals are integrated by a multiplication operator, defined in Eq (3). Since there are forty disparity tuned cells and five vergence tuned cells, we obtain 200 gain-modulated cells in MT. Distance information is estimated in LIP as a linear combination of the weighted outputs from MT. It has been generally assumed that humans accurately perceive geometrical distance based on disparity and vergence. However, people tend to overestimate close viewing distances and underestimate far distances in reality [24], even within the arm’s reaching space [26]. Therefore, we transformed geometrical distance into perceived distance (see S1 File for details). This perceived distance can be a teaching signal to train the network to gain the optimal weights. The delta rule, an iterative optimization technique, is used to find a set of optimal weights for the network.

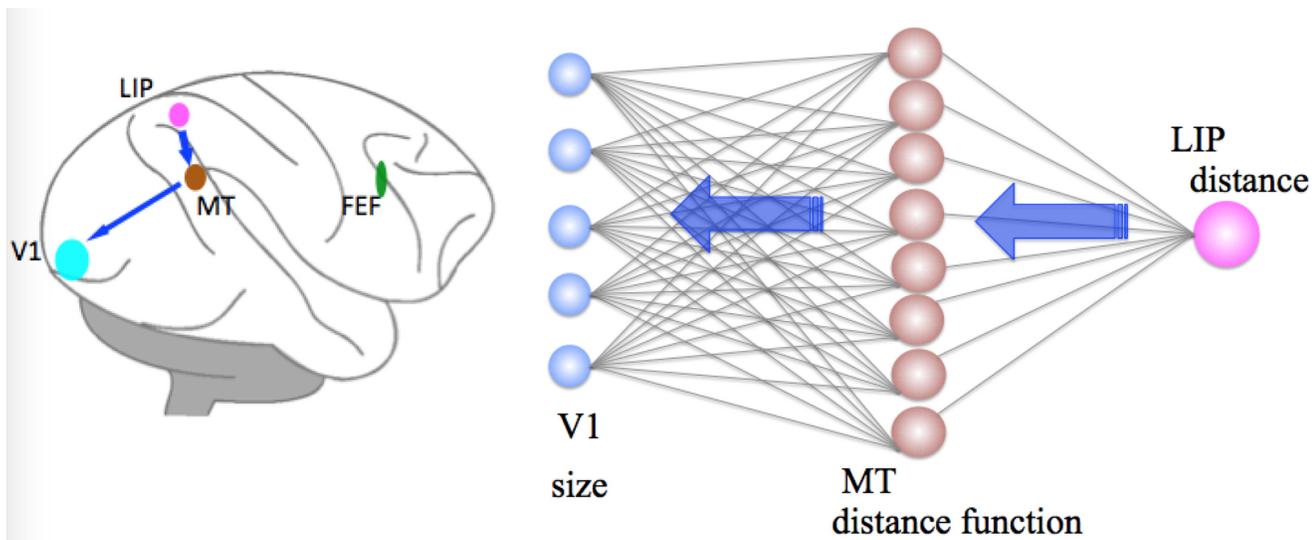
### Size Module

The size module consists of feedback connections from LIP to MT and further back to V1 (Fig 6). As we described above, the output of the distance module is a distance map estimated from a set of vergence and disparity input pairs. The distance information feeds back to the size selective cells in V1 through MT, where a distance scaling function is constructed. The distance scaling function modulates the response of size tuned cells in V1. This function can be used to obtain nearness, farness, and distance-independent cells identified in Dobbins et al. [10].

The size-selective cell,  $s_i$ , is defined by Gaussian:

$$s_i = e^{-\frac{(\beta - \beta_i)^2}{\sigma_i^2}} \tag{8}$$

where  $\beta_i$  is the peak of response for a given curve,  $\sigma_i$  describes the width of the tuning curve, and  $K$  indicates a distance scaling factor incorporated by MT cells.  $\beta$  is within  $[0^\circ, 3.2^\circ]$ . Because there are nearness, farness, and distance-independent cells, different distance scaling functions are set to accommodate the diversity in cell responses. For distance-independent

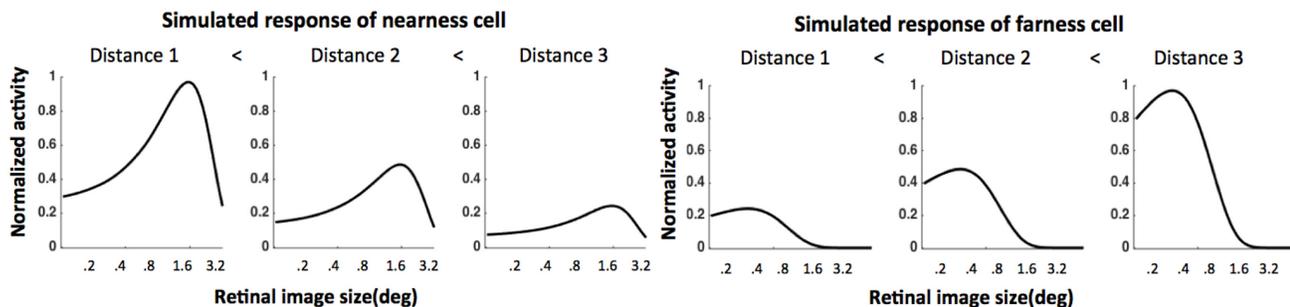


**Fig 6. Size module.** The distance information (see variable  $D$ , Eq (2)) in LIP feeds back to size tuned cells (see variable  $s$ , Eq (8)) in V1 through MT. size tuned cells in V1 are gain-modulated and have Gaussian-shaped tuning curves. In area MT, distance scaling functions are constructed. The activity of the size tuned cells is modulated by distance information through distance-dependent shunting equation (Eq (9)).

doi:10.1371/journal.pone.0129377.g006

cells, the distance scaling function equals a constant, i.e.,  $K = C$ , cell response does not change over distance. For farness cells,  $K = CD(\delta, \nu)$ , where  $D(\delta, \nu)$  is the distance information obtained from the distance module, cell response increases over distance. For nearness cells,  $K = D(\delta, \nu)$ , cell response decreases over distance. Fig 7 shows that the shape and the position of a simulated nearness/farness cell in response to stimuli with equal retinal image sizes are constant, but the amplitude of the response decrease/increase as the distance increases, just as shown in Fig 1a and 1b in Dobbins' study.

The model thus far has shown a distance-dependent property of size representation. In addition, in order to reflect size constancy in V1, cortical activity should be relatively constant when viewing an object (with constant physical size) at varying distances. Conversely, when viewing an object with constant angular size at varying distances, the spread of cortical activity in V1 should change [14]. That is to say, with a fixed angular size, the farther the object is, more eccentric parts of V1 get involved. Here, we used a similar stimuli paradigm as in



**Fig 7. Simulated tuning curves of nearness and farness cells.** The cell tuning curves in V1 are simulated by Gaussians. Similar to Fig 1a and 1b in Dobbins et al [10], the magnitudes of the responses of nearness (farness) cells decrease (increase) with viewing distance (Distance 1 < Distance 2 < Distance 3), but the shape and position of the peaks of the tuning curves are similar.

doi:10.1371/journal.pone.0129377.g007

Sperandio et al. to demonstrate the spread of cortical activity: on each trial, stimuli were presented twice (Sperandio et al. used an afterimage instead of a second stimulus presentation). The second stimulus presentation was at one of the three different viewing distances, provoking maximum activation of the corresponding three size tuned cells. In order to demonstrate the cortical activity spread, our model used distance-dependent shunting network equations [46] to simulate the dynamics of change in activity of the size tuned cells. The distance-dependent shunting equation involves basic membrane, or so-called shunting, equation coupled by distance-dependent interactions. Responses of cells provoked by stimulus inputs are regulated by these interactions, and therefore demonstrate distance-related properties. For example, topographic shifts of cell responses in a neural population may be observed in accordance with the spatial shifts of stimulus. In our model, size tuned cells in V1 obey the distance-dependent shunting equations, whereby stimulus input  $I_k$  at cell  $k$  is regulated by Gaussian defined excitatory and inhibitory connections among cells ( $s_{ki}$  and  $s'_{ki}$  where  $k \neq i$ ). It is critical to simulate the dynamics of cells' activity, from which we can see the temporal shifts in cell responses affected by distance-dependent interactions. The equation is defined as follows:

$$\frac{dS_i}{dt} = -B_1 S_i - (B_2 - S_i) \sum_{k=1}^n I_k s_{ki} - (B_3 + S_i) \sum_{k=1}^n I_k s'_{ki} \quad (9)$$

where  $S_i$  is the response of distance-dependent size tuned cell,  $i$ . In the early visual cortex, most feature-selective cells are organized topographically such that neighboring cells respond to similar quantities of a feature [47]. In our model, we employ the same orderly organization pattern for size-selective cells in V1, so neighboring cells are selective to similar angular sizes. Thus, the excitatory and inhibitory Gaussian kernels are defined by Eq (8). The decay rate of cell activity,  $B_1$ , was also regulated by the distance scaling function. In this simulation,  $B_1$  was scaled by  $CD$  ( $\delta, \nu$ ) to simulate the farness cell response. Parameters  $B_2$  and  $B_3$  are chosen to constrain the permitted values of  $S_i$ ,  $S_i > 0$ .

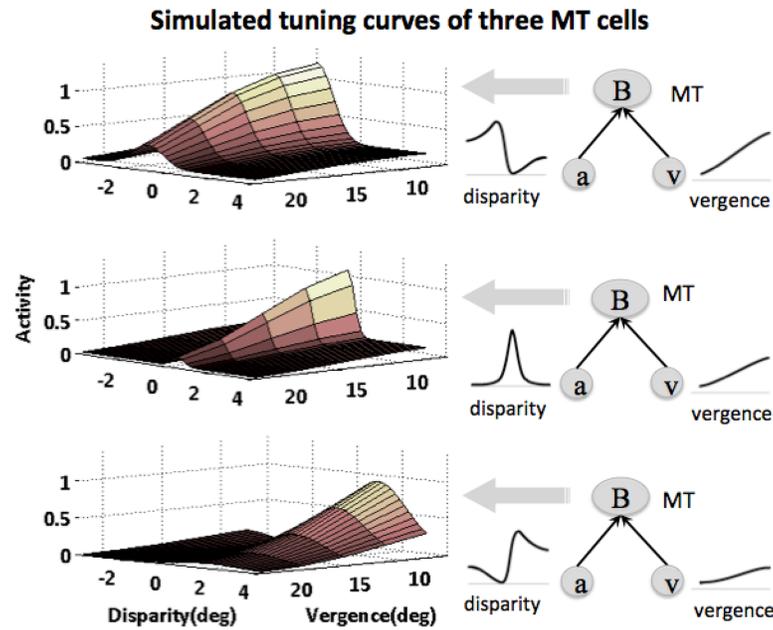
## Results

At the level of MT, disparity and vergence signals are integrated. Fig 8 shows four typical examples of MT cells depicted in the previous section (Eq (3)). Each panel illustrates the simulated activity of a single model MT cell in response to HD and vergence. As shown in the figure, the responses of disparity-selective cells were gated by vergence signals.

The results of LIP simulation are shown in Fig 9. The top panel depicts geometric distance as a function of vergence and disparity; the middle panel displays the approximate perceived distance; the bottom panel illustrates the approximations by a network of gain modulated units. This distance information output from the model's LIP then feeds back to MT and to V1 to modulate the activity of size tuned cells.

The cortical activity spread in V1 reported by Sperandio et al. [14] is demonstrated in Figs 10 and 11a. The perceived eccentricity differences are reflected by shifts in the distribution of activity across the surface of V1. Activation in response to perceptually larger stimuli (an afterimage) occurred in a more eccentric position in V1 compared to perceptually smaller stimuli with the same angular size.

Fig 11b shows the simulated cortical activity in V1. As specified in the Methods section, the stimulus was first present at the beginning of the time courses, then it was turned off. After a short interval, the stimulus was turned back on at one of the three different viewing distances, distance 1 < distance 2 < distance 3. Comparing the three panels on the right shows that at the beginning, the activity was strongest at the least eccentric region of interest (Eccentricity 1,



**Fig 8. Simulated tuning curves for MT cells in response to vergence and disparity.** MT cell responses to disparity is gated by vergence, Three examples of cell activity are shown, corresponding to disparity cells (a) tuned to crossed, uncrossed and zero disparity in V1, and vergence cells (v) in FEF with different tuning amplitudes.

doi:10.1371/journal.pone.0129377.g008

marked in red) for all three viewing distances tested, consistent with experimental data from Sperandio et al. [14]. When a stimulus with the same angular size (or an afterimage) was presented at a greater distance, corresponding activity was stronger in the more eccentric region of interest (ROI). For example, compare the bottom panel (distance 3) to the above two panels (distance 1 & distance 2), the ROI related to the largest eccentricity (Ecc 5, marked in purple) gradually increases its activity with viewing distance, and shows the greatest activity at the farthest viewing distance. In other words, the larger the stimuli appeared, the more eccentric the activation in V1.

### Quantitative comparison

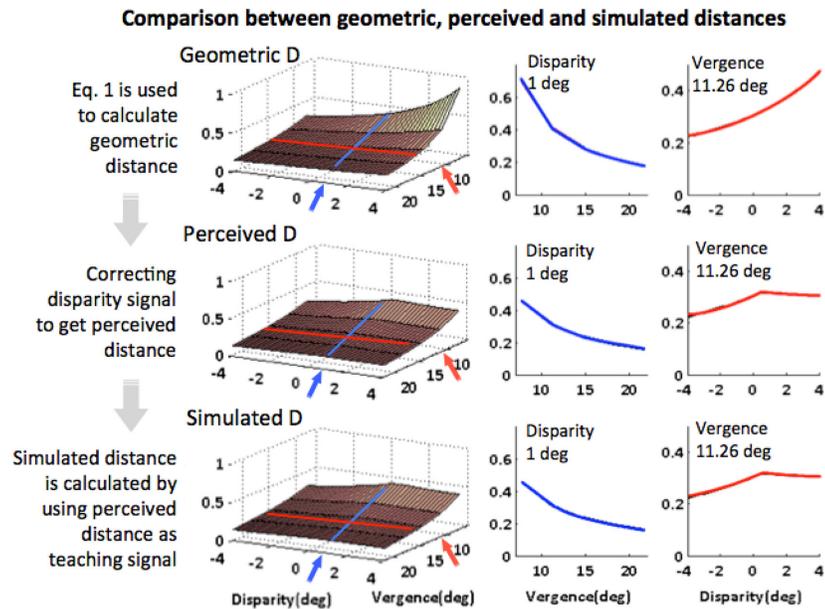
We used three measures to compare the quality of the simulation fits to the empirical data: 1) cosine of correlation angle (CCA); 2) difference in root mean square (dRMS); 3) Pearson product-moment correlation coefficient,  $r$ .

**Cosine of correlation angle (CCA).** The CCA is defined as the inner product of the simulation and the BOLD signal curve vectors divided by their norms (see Eq (10)). If the two vectors are parallel with each other, CCA equals 1. The closer it gets to 1, the better the fit is.

$$CCA = \frac{\langle F, S \rangle}{\|F\| \cdot \|S\|} \quad (10)$$

where  $F$  is the vector of the BOLD signal change, and  $S$  is the vector of the model simulation.

**Difference in root mean square (dRMS).** Root mean square is another metric that can be extracted directly from the time-domain signal. It is defined as the square root of the average

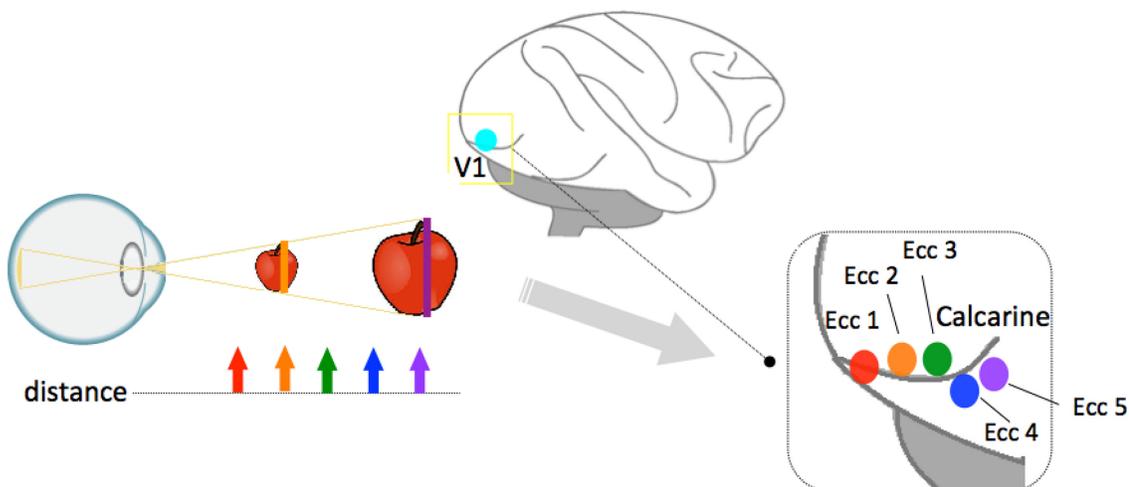


**Fig 9. Distance cell response as a function of vergence and disparity in LIP.** The top left panel: geometric distance; middle left: approximate perceived distance, calculated by correcting disparity signal in geometric distance (see text and [S1 File](#)); bottom left: distance cell response in LIP, simulated by training a network of gain-modulated units of disparity and vergence cells using perceived distance as teaching signal. The right panels show two examples of distance approximation as a function of vergence/disparity, taken from the cross sections (indicated by the red arrows) of the distance response on the left. Y-axes labels for the right panels are the same as of the left ones.

doi:10.1371/journal.pone.0129377.g009

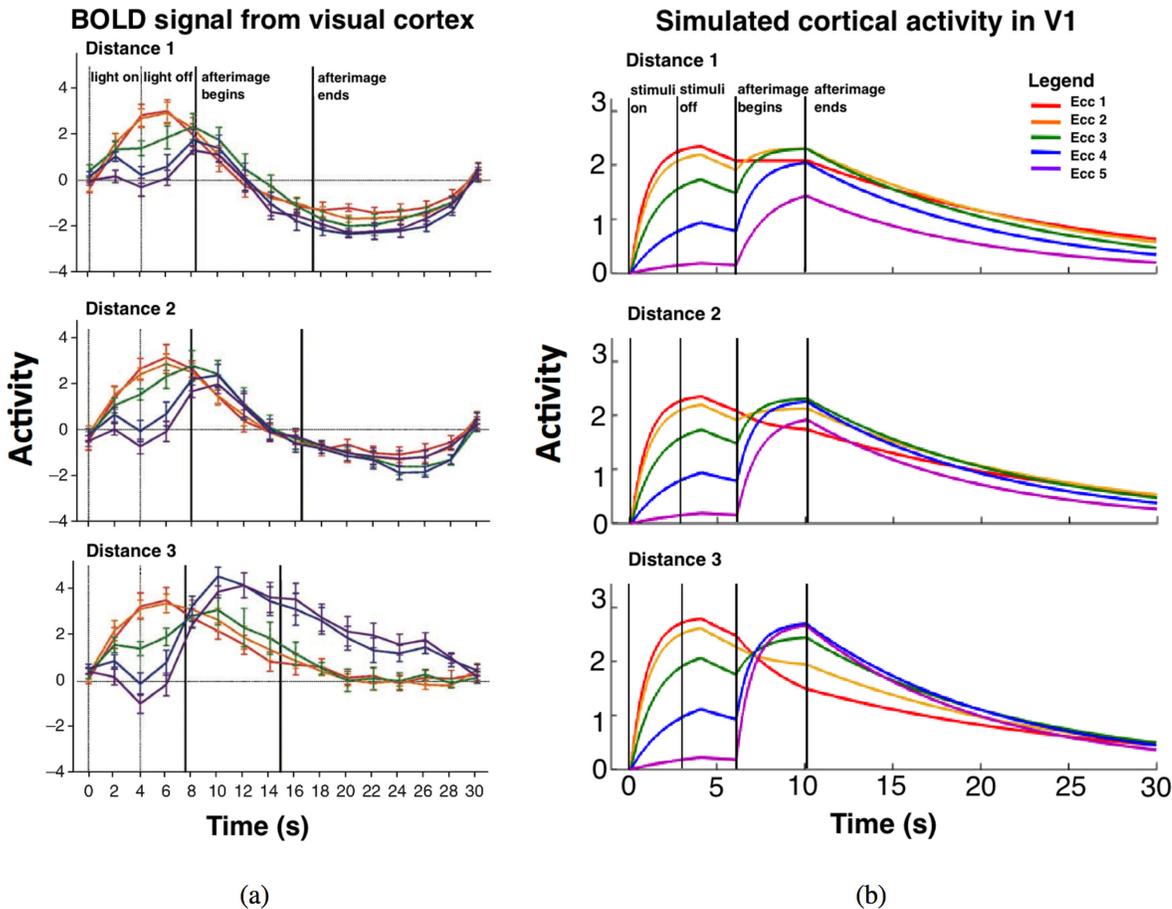
squared value of the signal and can also be called the normalized energy of the signal:

$$RMS = \sqrt{\frac{1}{n} \sum_{i=0}^n S_i^2} \tag{11}$$



**Fig 10. Cortical activity in V1 modulated by perceived size.** When a stimulus is presented at a close distance, the activity is strongest in the smaller eccentricity along the calcarine sulcus; when the stimulus with the same angular size (or an afterimage) is presented farther, the activity is stronger in the more eccentric areas. In other words, the bigger the stimuli perceived, the more eccentric the activation in V1. Different colors mark the different eccentric activation corresponding to different viewing distance, e.g., red marks the smallest eccentric activation and purple marks a largest eccentric activation in V1.

doi:10.1371/journal.pone.0129377.g010



**Fig 11. Comparison between experimental data and simulated cortical activity in V1.** Panel (a) shows the BOLD signal change recorded from human visual cortex V1 [14]; Panel (b) shows the simulated cortical activity in V1 by our model, which resembles the experimental data on the left. At the beginning of the time courses, when the stimulus is presented at the nearest distance, the activity was strongest in the smaller eccentricity ROIs (marked in red and orange); when the stimulus with the same angular size is presented at a farther distance (one of the three viewing distances tested), the activity is stronger in the more eccentric ROIs (purple and blue) as viewing distance increased. Distance 1 < Distance 2 < Distance 3. Refer to Fig 10 for color marking. Panel (a) is adapted from Sperandio et al. [14].

doi:10.1371/journal.pone.0129377.g011

The dRMS is defined as

$$dRMS = \frac{RMS(F - S)}{RMS(F) + RMS(S)} \tag{12}$$

which shows how much the curves are off from each other, normalized by their own values. The closer it gets to 0, the less the difference is.

**Pearson product-moment correlation coefficient, r.** Pearson's r is defined as

$$r = \frac{\cos(F, S)}{\sqrt{\text{var}(F)\text{var}(S)}} \tag{13}$$

Quantitative comparison shows that 1) the CCA equals .903, with 95% confidence interval (CI) of [.806, 1.000]; 2) the dRMS equals .288, with 95% CI of [.098, .479]; 3) Pearson's r equals .622,  $p < .001$ .

## Discussion

We propose a neural model that achieves distance-dependent size tuning in V1. The model consists of bidirectional connections from V1 and FEF to LIP through MT. A subset of disparity-selective cells in V1 and vergence-selective cells in FEF contribute to the creation of a 3D spatial map in LIP through feedforward interareal connections via MT. The distance information from the 3D representation feeds back to a subset of size tuned cells in V1, modulating their responses and the spread of cortical activity. Our simulation of this model is consistent with neurophysiological findings [10, 14], which suggests that distance-dependent size representation in V1 is a manifestation of size constancy mechanisms.

Under normal viewing conditions, size constancy of humans is nearly perfect for distances under 100 feet [48–50]. Researchers have proposed that this prominent perceptual ability can be achieved by scaling retinal size to compensate for viewing distance. In accordance with a large body of psychophysical evidence [1–4], Dobbins et al.'s neurophysiological findings support this hypothesis, that the magnitudes of size tuned cells' responses in V1, V2 and V4 are modulated by viewing distance [10]. They suggested that distance-dependent modulation of cell's responses is possibly a common property of neurons along the ventral visual pathway, from V1 to IT. Apart from size, other features may share this property. For example, disparity tuned cells in V1 show similar activity patterns in response to change in viewing distance [22]. In addition, distance and angle of gaze modulate neural responses in parietal cortex as well as on the dorsal pathway from V1 to parietal cortex [16]. These results suggest that spatial modulation exists in both the dorsal and ventral visual cortical streams, and appears to be a fundamental attribute of the visual cortex.

In order to realize spatial modulation, an estimate of distance is necessary for visual perception so that distance scaling can be applied to all related visuocortical areas. A few neural models have been proposed to address the distance estimation problem. In particular, Pouget et al. [21] proposed a neural network showing that cortical representation of egocentric distance could be computed by combining the outputs from gain-modulated Gaussian units and sigmoids, simulating the responses of disparity- and vergence- selective cells, respectively. Although Pouget et al. did not tackle the problem of size constancy, they mentioned that the neural mechanisms underlying perceptual constancies might adopt a strategy similar to the strategy of distance estimation (gating the activity of disparity-selective neurons by vergence signals). In other words, the activity of certain feature-selective neurons was gated by distance signals to achieve constancy. Other than their study, a few neural models were proposed to address similar topics [51, 52]. But till present, no previous modeling work has explicitly demonstrated the neural mechanism of size constancy.

Our model adopts a similar method, as proposed in Pouget et al., to estimate distance. Distance information modulates the neural activity of size tuned cells, providing distance-dependent size tuning reported by Dobbins et al. [10]. In addition, our model simulates recent findings that demonstrate associations between the spatial extent of activation in human primary visual cortex and an object's perceived size [12–14]. Activity in V1 becomes more eccentric as the perceived size of a real stimulus, or an afterimage, increases, even though the size of the retinal image remains the same. The retinal signals reaching V1 are modulated by distance information in a fashion that reflects the implementation of size-constancy mechanisms. It is suggested that the spread in cortical location can hardly be explained by local interactions between the stimuli and the 3D scene (e.g., local contrast), because the local interaction would have increased or decreased the overall magnitude of neural activity, rather than induced the cortical activity spread. In addition, attention allocation is required to perform size constancy, which rules out stimulus-based explanations for the spatial spread in activity [13]. Feedback

from higher visual areas would seem to make an important contribution to this effect. Specifically, extracting distance information from depth cues, such as linear perspective and texture cues, requires that the information be integrated over a large area with large receptive fields comparable in size to those found in higher visual areas.

In our proposed model, depth cues are integrated along the dorsal pathway leading to LIP. Empirical studies show that LIP may provide egocentric distance representation [16, 38–40]. We further propose that integration of vergence and disparity takes place in MT. Although Chowdhury et al. [53] showed that vertical disparity cues to viewing distance did not modulate MT horizontal disparity tuning, currently no study has directly examined whether the responses of disparity selective cells in MT are gated by vergence signals. Given MT's anatomical connectivity with LIP, V1, and FEF [35–37], we suggest that it is an appropriate candidate for integrating disparity and vergence. In addition, there are studies showing that the role of MT is not limited to processing moving objects. For example, many MT neurons exhibit robust disparity tuning in response to stationary random-dot patterns [54], and are involved in representing implied motion from static images [55, 56]. On the other hand, in the ventral stream, a possible integration area is V3v. A recent study showed that left V3v may be involved in integrating retinal size and distance information [57]. But to our knowledge, currently no area found along the ventral stream directly codes distance information. Therefore one possibility is that V3v receives distance feedback from LIP via MT and V3A. However, these proposed functions of the brain areas require further investigation and confirmation.

Fig 11b simulates the spread of cortical activity in V1 reported by Sperandio et al. [14]. Quantitative comparison shows that the simulation curve fits the BOLD signal well. Although the overall shape of the simulation and the empirical data matches, there are small differences. The simulated curves have two distinctive phases while the empirical data does not. This can be attributed to the fact that BOLD signal change reflects the activity of a population of neurons, while simulated data reflects activity from a single neuron. For the empirical data, Eccentricity 1–3 showed more activity during the light-on duration (and consequently followed by the light-off duration) than Eccentricity 4&5, because the annuli sizes of the ROI localizers for Ecc 1–3 were close to the light size used to induce the afterimage (in terms of retinal projection). The lack of high amplitude in BOLD response in Ecc 4&5 was followed by a gradual amplitude elevation within the light-off period and continued further during afterimage interval. This indicates that there are time-consuming neural processes going-on to build up and feedback the higher eccentricity parts of V1. These all give the appearance to the empirical data as if in Ecc 4&5 there are no two distinctive phases. Although the positions of peaks of the BOLD signal curves almost stay the same through Distance 1 to 3, the higher-eccentricity parts of V1 in response to the afterimages viewed at larger distances (and hence perceptually larger) seemed to have delayed peaks (as if there is a right peak shift). This delayed BOLD amplitude elevation could also be due to the time-consuming neural processes in bringing activities back to higher-eccentricity parts in V1. In our simulation, there was a surge of activities at the start of the afterimage, and this rebound of activity gave the bi-phasic appearance to the simulated curves. This is that the simulation shows cell response, while the fMRI data is "low-pass filtered" in time domain and hence the rebound of activity could be smoothed out. Besides this, other aspects are consistent with the fMRI data.

The simulation results were achieved by first modulating the responses of size tuned cells in V1 with distance scaling, which is derived in MT through feedback from LIP. Second, our model of V1 is topographically organized by size selectivity; this allows spreading of cortical activation when using distance-dependent shunting equations to simulate the dynamics of cell activity. The result supports that the topographic representation in V1 is dependent on the egocentric distance estimation. The model predicts that the neurons classical receptive fields shift

to more eccentric cortical locus. As the perceived distance increases, more neurons are activated, therefore shows a spatial spread of cortical activity. Moreover, in order to simulate the response pattern seen in [Fig 11a](#), distance feedback needed to modulate the activities of V1 farness cells in the simulation. Therefore, the model predicts that the type of size-selective cell (see [Fig 7](#)) affects the size of neurons population receptive fields. For example, when an object gets farther away, the farness cells get more involved and the size of their population RF increases. In addition, an important implication of the model is that the spatial spread of cortical activity may be found in other brain areas where cells show topographic organization, and where their activity (possibly in response to features other than size) is affected by distance. The distance information, once extracted, can be used to rescale retinotopic representations in other visual areas. It is possible that the cortical activity spread is a neural marker coding for changes in viewing distance. This prediction can be tested in neurophysiological experiments that may further shed light on the neural mechanisms of distance-related feature perception. For example, neurophysiological methods could assess whether cells selective for curvature or contrast demonstrate similar response properties (like dependent spreading of cortical activation). It has been shown that perceptual constancies may be intercorrelated, e.g. size and contrast constancy, and it is possible that these phenomena share a common scaling factor [\[8, 9\]](#).

In our model, disparity and vergence are the only two depth cues used. In addition to these two cues, numerous psychophysical studies have shown that the vertical disparity (VD) can provide viewing distance information independent of vergence [\[58–60\]](#). Rogers & Bradshaw [\[18\]](#) showed that it can be used to judge absolute distance and disparity scaling. Durand et al. [\[61, 62\]](#) found that VD encoding shared a common neural process with the HD encoding, with narrower widths of the tuning curves and narrower range for VD. The range extended to a larger eccentric angular scale in parafoveal V1, while centered on a VD value of 0° in foveal V1 [\[63\]](#). This discrepancy might due to the fact that VDs do not naturally occur in central vision, typically for objects smaller than 20° [\[18, 64\]](#). Although VD is unlikely to be effective in specific experimental situations that our model addressed, it can be easily integrated into our model. Since egocentric distance is usually constrained in laboratory settings, it is reasonable to assume that disparity and vergence are the dominant depth cues in our simulation.

In the real world, distance can be infinitely far where disparity and vergence cues become less effective if not absolutely ineffective. Many other depth cues may be used, like pictorial and motion-related depth cues. Although not conventionally deemed as depth cues, looming or optic flow, which provides information about an observer's self-motion, could be integrated to estimate distance travelled as well. Therefore, other depth cues may gradually take over the dominant role and contribute significantly in distance estimation. For instance, one computational model has tackled the problem of time to contact by integrating disparity and looming cues, weighted by stimulus size [\[53\]](#). Estimating time to contact shares some common factors with estimating distance, like involving integration of different depth cues. In principle, the more coherent depth cues can be utilized, the more accurate the distance estimate. It is therefore important for a neural model to have compatibility. Our model can accommodate such situations, e.g., vertical disparity, optic flow, etc. Vertical disparity can be included by extension in the input representation. Besides disparity and vergence, it can be a third variable used in the basis function set to estimate distance. As for optic flow, it is known that neurons in MST are selective for its direction and speed [\[65\]](#). With cortical connections from MST to MT [\[66\]](#), depth information extracted from optic flow can feed MT, where it integrates other depth cues.

The visual areas used in our model are by no means exhaustive, and it is extremely likely that many other visual areas are involved. Further research may shed light on other possible brain areas contributing to constancy mechanisms. Investigating these questions may clarify

how distance information can be estimated or used to modulate cell activity, and therefore improve our understanding of perceptual constancy phenomena.

## Conclusions

Recent research has shown that the activity of size tuned cells in V1 is possibly regulated by feedback of distance information from other brain areas, demonstrating size constancy. We propose a physiologically based neural model that simulates activity of distance-dependent size tuned cells in V1 [10, 14]. The model suggests that many brain areas may share a common set of distance scaling functions, which is essential for three-dimensional spatial coding in visual processing. In addition, we speculate that the topographic spread of cortical activity may be found in other brain areas whose activity (possibly in response to features other than size, like contrast) is modulated by distance information.

## Supporting Information

### S1 File. Training algorithm.

(DOCX)

### S1 Table. Variables and their definitions.

(DOCX)

## Acknowledgments

We thank Celia Gagliardi and Jeffrey Nador for help with proofreading the manuscript and the Department of Psychology at Northeastern University, in particular Joanne Miller for their support during the progress and final stages of the research.

This work has been supported in part by the Center of Excellence for Learning in Education, Science and Technology (CELEST, NSF SBE-0354378), Office of Naval Research (ONR N00014-11-1-0535), Air Force Office of Scientific Research (AFOSR, FA9550-12-1-0436), and the Fundamental Research Funds for the Central Universities (14wkpy20). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author Contributions

Conceived and designed the experiments: JQ AY. Performed the experiments: JQ. Analyzed the data: JQ. Contributed reagents/materials/analysis tools: JQ AY. Wrote the paper: JQ AY.

## References

1. Boring EG (1940). Size constancy and Emmert's law. *The American Journal of Psychology*, 53(2), 293–295.
2. Epstein W (1963). Attitudes of judgment and the size-distance invariance hypothesis. *Journal of Experimental Psychology*, 66(1), 78–83.
3. Epstein W, Park J, Casey A (1961). The current status of the size-distance hypotheses. *Psychological Bulletin*, 58(6), 491–514.
4. Kaufman L, Kaufman JH, Noble R, Edlund S, Bai S, King T (2006). Perceptual distance and the constancy of size and stereoscopic depth. *Spatial vision*, 19(5), 439–457. PMID: [17131650](https://pubmed.ncbi.nlm.nih.gov/17131650/)
5. Dees JW (1966). Moon illusion and size-distance invariance: An explanation based upon an experimental artifact. *Perceptual and Motor Skills*, 23(2), 629–630.
6. Ross HE (1967). Water, fog and the size—distance invariance hypothesis. *British Journal of Psychology*, 58(3–4), 301–313.

7. Ross HE (2000). Cleomedes (c. 1st century AD) on the celestial illusion, atmospheric enlargement, and size-distance invariance. *PERCEPTION LONDON*, 29(7), 863–872. PMID: [11064807](#)
8. Qian J, Petrov Y (2012). Startrek illusion—general object constancy phenomenon? *Journal of Vision*, 12(2).
9. Qian J, Petrov Y (2013). Depth perception in the framework of general object constancy. *Journal of vision*, 13(11).
10. Dobbins AC, Jeo RM, Fiser J, Allman JM (1998). Distance modulation of neural activity in the visual cortex. *Science*, 281(5376), 552–555. PMID: [9677196](#)
11. Ito M, Tamura H, Fujita I, Tanaka K (1995). Size and position invariance of neuronal responses in monkey inferotemporal cortex. *Journal of neurophysiology*, 73(1), 218–226. PMID: [7714567](#)
12. Murray SO, Boyaci H, Kersten D (2006). The representation of perceived angular size in human primary visual cortex. *Nature neuroscience*, 9(3), 429–434. PMID: [16462737](#)
13. Fang F, Boyaci H, Kersten D, Murray SO (2008). Attention-dependent representation of a size illusion in human v1. *Current Biology*, 18(21), 1707–1712. doi: [10.1016/j.cub.2008.09.025](#) PMID: [18993076](#)
14. Sperandio I, Chouinard PA, Goodale MA (2012). Retinotopic activity in V1 reflects the perceived and not the retinal size of an afterimage. *Nature Neuroscience*, 15(4), 540–542. doi: [10.1038/nn.3069](#) PMID: [22406550](#)
15. Liu Q, Wu Y, Yang Q, Campos JL, Zhang Q, Sun HJ (2009). Neural correlates of size illusions: an event-related potential study. *NeuroReport*, 20(8), 809–814. doi: [10.1097/WNR.0b013e32832832be7c0](#) PMID: [19384256](#)
16. Gnadt JW, Mays LE (1995). Neurons in monkey parietal area lip are tuned for eye-movement parameters in three-dimensional space. *Journal of neurophysiology*, 73(1), 280–297. PMID: [7714572](#)
17. Cumming B, DeAngelis G (2001). The physiology of stereopsis. *Annual review of neuroscience*, 24(1), 203–238.
18. Rogers BJ, Bradshaw MF (1993). Vertical disparities, differential perspective and binocular stereopsis. *Nature*, 361(6409), 253–255. PMID: [8423851](#)
19. Richards W (1971b). Size-distance transformations. *Pattern Recognition in Biological and Technical Systems* (pp. 276–287). Springer.
20. Richards W, Kaye MG (1974). Local versus global stereopsis: two mechanisms? *Vision Research*, 14(12), 1345–1347. PMID: [4446365](#)
21. Pouget A, Sejnowski TJ, et al. (1994). A neural model of the cortical representation of egocentric distance. *Cerebral Cortex*, 4, 314–314. PMID: [8075535](#)
22. Trotter Y, Celebrini S, Stricanne B, Thorpe S, Imbert M (1992). Modulation of neural stereoscopic processing in primate area v1 by the viewing distance. *Science*, 257(5074), 1279–1281. PMID: [1519066](#)
23. Trotter Y, Celebrini S, Stricanne B, Thorpe S, Imbert M (1996). Neural processing of stereopsis as a function of viewing distance in primate visual cortical area V1. *Journal of Neurophysiology*, 76(5), 2872–2885. PMID: [8930240](#)
24. Foley J (1980). Binocular distance perception. *Psychological Review*, 87(5), 411. PMID: [7413886](#)
25. Priot AE, Neveu P, Sillan O, Plantier J, Roumes C, Prablanc C (2012). How perceived egocentric distance varies with changes in tonic vergence. *Experimental Brain Research*, 219, 457–465. doi: [10.1007/s00221-012-3097-5](#) PMID: [22623089](#)
26. Viguier A, Clement G, Trotter Y (2001). Distance perception within near visual space. *Perception*, 30(1), 115–24. PMID: [11257974](#)
27. Poggio G, Fischer B, et al. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology*, 40(6), 1392–1405. PMID: [411898](#)
28. Poggio GF, Gonzalez F, Krause F (1988). Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity. *The Journal of Neuroscience*, 8(12), 4531–4550. PMID: [3199191](#)
29. Poggio GF, Motter BC, Squatrito S, Trotter Y (1985). Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Research*, 25(3), 397–406. PMID: [4024459](#)
30. Gonzalez F, Perez R (1998). Neural mechanisms underlying stereoscopic vision. *Progress in Neurobiology*, 55(3), 191. PMID: [9643554](#)
31. Mays LE, Porter JD, Gamlin P, Tello CA (1986). Neural control of vergence eye movements: neurons encoding vergence velocity. *Journal of Neurophysiology*, 56(4), 1007–1021. PMID: [3783225](#)
32. Gamlin PD, Mays LE (1992). Dynamic properties of medial rectus motoneurons during vergence eye movements. *Journal of Neurophysiology*, 67(1), 64–74. PMID: [1552323](#)

33. Gamlin PD, Yoon K (2000). An area for vergence eye movement in primate frontal cortex. *Nature*, 407(6807), 1003–1007. PMID: [11069179](#)
34. Gamlin PD (2002). Neural mechanisms for the control of vergence eye movements. *Annals of the New York Academy of Sciences*, 956(1), 264–272.
35. Hutchison RM, Gallivan JP, Culham JC, Gati JS, Menon RS, Everling S (2012). Functional connectivity of the frontal eye fields in humans and macaque monkeys investigated with resting-state fMRI. *Journal of Neurophysiology*, 107(9), 2463–2474. doi: [10.1152/jn.00891.2011](#) PMID: [22298826](#)
36. Girard P, Hupe JM, Bullier J (2001). Feedforward and feedback connections between areas v1 and v2 of the monkey have similar rapid conduction velocities. *Journal of Neurophysiology*, 85(3), 1328–1331. PMID: [11248002](#)
37. Nowak LG, Bullier J, et al. (1998). The timing of information transfer in the visual system. *Cerebral Cortex*, 12, 205–241.
38. Andersen RA, Essick GK, Siegel RM (1985). Encoding of spatial location by posterior parietal neurons. *Science*, 230(4724), 456–458. PMID: [4048942](#)
39. Blatt GJ, Andersen RA, Stoner GR (1990). Visual receptive field organization and cortico-cortical connections of the lateral intraparietal area (area LIP) in the macaque. *Journal of Comparative Neurology*, 299(4), 421–445. PMID: [2243159](#)
40. Genovesio A, Ferraina S (2004). Integration of retinal disparity and fixation-distance related signals toward an egocentric coding of distance in the posterior parietal cortex of primates. *Journal of Neurophysiology*, 91(6), 2670–2684. PMID: [14960558](#)
41. Ninomiya T, Sawamura H, Inoue K, Takada M (2012). Segregated pathways carrying frontally derived top-down signals to visual areas MT and V4 in macaques. *The Journal of Neuroscience*, 32(20), 6851–6858. doi: [10.1523/JNEUROSCI.6295-11.2012](#) PMID: [22593054](#)
42. Jones J, Palmer L (1987). An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58(6), 1233–1258. PMID: [3437332](#)
43. Marcelja S (1980). Mathematical description of the responses of simple cortical cells. *Journal of the Optical Society of America*, 70(11), 1297–1300. PMID: [7463179](#)
44. Daugman JG (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A-optics image science and vision*, 2(7), 1160–1169.
45. Lehky S, Sejnowski T (1990). Neural network model of visual cortex for determining surface curvature from images of shaded surfaces. *Proceedings of the Royal Society of London. B. Biological Sciences*, 240(1298), 251–278.
46. Grossberg S (1983). The quantized geometry of visual space: The coherent computation of depth, form, and lightness. *Behavioral and Brain Sciences*, 6(625–57),
47. Bonhoeffer T, Grinvald A (1991 Oct 3). Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. *Nature*, 353(6343), 429–431. PMID: [1896085](#)
48. Holway AH, Boring EG (1941). Determinants of apparent visual size with distance variant. *The American Journal of Psychology*, 54(1), 21–37.
49. Gilinsky AS (1955). The effect of attitude upon the perception of size. *The American Journal of Psychology*, 68(2), 173–192. PMID: [14376680](#)
50. Ungerleider LG, Ganz L, Pribram KH (1977). Size constancy in rhesus monkeys: Effects of pulvinar, prestriate, and inferotemporal lesions. *Experimental Brain Research*, 27(3–4).
51. Anderson CH, Van Essen DC (1987). Shifter circuits: a computational strategy for dynamic aspects of visual processing. *Proceedings of the National Academy of Sciences*, 84(17), 6297–6301.
52. Rushton SK, Wann JP (1999). Weighted combination of size and disparity: a computational model for timing a ball catch. *Nature neuroscience*, 2(2), 186–190. PMID: [10195204](#)
53. Chowdhury SA, Christiansen DL, Morgan ML, DeAngelis GC (2008). Effect of vertical disparities on depth representation in macaque monkeys: MT physiology and behavior. *Journal of Neurophysiology*, 99(2), 876–887. PMID: [18077666](#)
54. DeAngelis GC, Cumming BG, Newsome WT (1998). Cortical area MT and the perception of stereoscopic depth. *Nature*, 394(6694), 677–680. PMID: [9716130](#)
55. Kourtzi Z, Bülthoff HH, Erb M, Grodd W (2002). Object-selective responses in the human motion area MT/MST. *Nature Neuroscience*, 5(1), 17–18. PMID: [11740503](#)
56. Kourtzi Z, Kanwisher N (2000). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience*, 12(1), 48–55. PMID: [10769305](#)

57. Weidner R, Plewan T, Chen Q, Buchner A, Weiss PH, Fink GR (2014). The moon illusion and size-distance scaling—evidence for shared neural patterns. *Journal of Cognitive Neuroscience*, 26(8), 1871–1882. doi: [10.1162/jocn\\_a.00590](https://doi.org/10.1162/jocn_a.00590) PMID: [24564459](https://pubmed.ncbi.nlm.nih.gov/24564459/)
58. Mayhew JE, Longuet-Higgins HC (1982). A computational model of binocular depth perception. *Nature*, 297, 376–378. PMID: [7078648](https://pubmed.ncbi.nlm.nih.gov/7078648/)
59. Bishop P (1989). Vertical disparity, egocentric distance and stereoscopic depth constancy: a new interpretation. *Proceedings of the Royal Society of London. B. Biological Sciences*, 237(1289), 445–469.
60. Bishop P (1994). Size constancy, depth constancy and vertical disparities: a further quantitative interpretation. *Biological Cybernetics*, 71(1), 37–47. PMID: [8054385](https://pubmed.ncbi.nlm.nih.gov/8054385/)
61. Durand JB, Celebrini S, Trotter Y (2007). Neural bases of stereopsis across visual field of the alert macaque monkey. *Cereb Cortex*, 17(6), 1260–73. PMID: [16908495](https://pubmed.ncbi.nlm.nih.gov/16908495/)
62. Durand JB, Zhu S, Celebrini S, Trotter Y (2002). Neurons in parafoveal areas V1 and V2 encode vertical and horizontal disparities. *J Neurophysiol*, 88(5), 2874–9. PMID: [12424321](https://pubmed.ncbi.nlm.nih.gov/12424321/)
63. Gonzalez F, Relova J, Perez R, Acuña C, Alonso JM (1993). Cell responses to vertical and horizontal retinal disparities in the monkey visual cortex. *Neuroscience Letters*, 160(2), 167–70. PMID: [8247349](https://pubmed.ncbi.nlm.nih.gov/8247349/)
64. Cumming BG, Johnston EB, Parker AJ (1991). Vertical disparities and perception of three-dimensional shape. *Nature*, 349, 411–413. PMID: [1992341](https://pubmed.ncbi.nlm.nih.gov/1992341/)
65. Duffy CJ, Wurtz RH (1991). Sensitivity of mst neurons to optic flow stimuli. i. a continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology*, 65(6), 1329–1345. PMID: [1875243](https://pubmed.ncbi.nlm.nih.gov/1875243/)
66. Ungerleider LG, Desimone R (1986). Cortical connections of visual area MT in the macaque. *Journal of Comparative Neurology*, 248(2), 190–222. PMID: [3722458](https://pubmed.ncbi.nlm.nih.gov/3722458/)