Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A New Discovery on Visual Information Dynamic Changes from Retina to V2

Haixin Zhong
East China University of Science and Technology
Rubin Wang (rbwang@163.com)
East China University of Science and Technology https://orcid.org/0000-0003-4110-2022

Research Article

Keywords: V2 encoding, dynamic degradation, convolution calculation, corner information, synaptic plasticity

Posted Date: April 6th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-354449/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Version of Record: A version of this preprint was published at Nonlinear Dynamics on August 5th, 2021.
See the published version at https://doi.org/10.1007/s11071-021-06648-0.

A new discovery on visual information dynamic changes from retina to V2

Haixin Zhong, Rubin Wang

Institute for Cognitive Neurodynamics, East China University of Science and Technology, 130 Meilong Road, Shanghai, People's Republic of China

Corresponding author Email: rbwang@163.com

Abstract: The information processing mechanisms of the visual nervous system remain to be unsolved scientific issues in neuroscience field, owing to a lack of unified and widely accepted theory for explanation. It has been well documented that approximately 80% of the rich and complicated perceptual information from the real world is transmitted to the visual cortex, only a small fraction of visual information reaches the V1 area. This, nevertheless, does not affect our visual perception. Furthermore, how neurons in V2 encode such a small amount of visual information has yet to be addressed. To this end, the current paper establishes a visual network model for retina-LGN-V1-V2 and quantitatively accounts for that response to the scarcity of visual information and encoding rules, based on the principle of neural mapping from V1 to V2. The results demonstrate that the visual information has a small degree of dynamic degradation when it is mapped from V1 to V2, during which there is a convolution calculation occurring. Therefore, visual information dynamic degradation mainly manifests itself along the pathway of the retina to V1, rather than V1 to V2. The slight changes in the visual information are attributable to the fact that the receptive fields (RFs) of V2 cannot further extract the image features. Meanwhile, despite the scarcity of visual information mapped from the retina, the RFs of V2 can still accurately respond to and encode "corner" information, due to the effects of synaptic plasticity, of which function is not existed in V1. This is a new discovery that has never been noticed before. To sum up, the coding of the "contour" feature (edge and corner) is achieved in the pathway of retina-LGN-V1-V2.

Keywords: V2 encoding, dynamic degradation, convolution calculation, corner information, synaptic plasticity

Abbreviations

RF: Receptive field

CV: Computer vision

LGN: Lateral geniculate nucleus

RMM: Recurrent motion model

STDP: Spike timing-dependent plasticity

LTP: Long-term potentiation

LTD: Long-term depression

PMVICV2: Predictive model for visual information changes in V2

DOG: Difference of two Gaussians

1 Introduction

The “Brain Projects” have been widely implemented throughout the world in recent years, such as those in China (Poo et al., 2016), the U.S.A (Bargmann & Newsome, 2014), Europe (Amunts et al., 2016), and Japan (Okano, Miyawaki, & Kasai, 2015). Such phenomenon has accordingly contributed to the burgeoning of research on visual information processing mechanisms in academic fields such as cognitive neuroscience and computer vision (CV) (Cox & Dean, 2014; Kendall & Kumar, 2020; Marblestone, Wayne, & Kording, 2016; Yu et al., 2020). Considered as the perfect image information processing system, visual system of human beings can quickly recognize such objective information as position, size, shape, color, orientation with substantial advantages in stability, robustness, efficiency and simplicity (Zhao, Zou, Jin, Yao, & Li, 2010). For that reason, scholars from fields of cognitive neurobiology, computational neuroscience, and CV have shown growing interest in examining the neural information processing mechanisms of the visual nervous system (Khan, Meyer, Konik, & Bouakaz, 2012; Oprea, Pack, & Khadra, 2020; Panetta, Gao, & Agaian, 2015; Raichle, 2010; Riley & Davies, 2020). Indeed, research on visual information processing mechanisms has kept accelerating as biological techniques continue to evolve over the past few decades. (Xu et al., 2017). In 1962, Wiesel and Hubel’s experimental research findings on cat’s V1 illustrated the correlation between the RFs of the lateral geniculate nucleus (LGN) and RFs of V1, which significantly advanced research in the field of biological vision (Hubel & Wiesel, 1962). In 1971, Dubner and Zeki studied the characteristics of the orientation selectivity of cells in V5, initially revealing that the MT area belongs to the central region of motion perception (Dubner & Zeki, 1971). In 1994, Ungerleider and Haxby proposed the theory of the dorsal pathway and ventral pathway, providing a physiological basis for the visual system to process motion and static information (Ungerleider & Haxby, 1994). In 2002, Riesenhuber and Poggio discovered mutual projection and interaction between the dorsal and ventral pathways. Building on this synergistic effect, the researchers further investigated perceptions under the influence of visual stimuli (Riesenhuber & Poggio, 2002). In the same year, Yifeng Zhou and Tiande Shou (2002) revealed that the orientation sensitivity of LGN cells could experience changes due to visual cortex feedback. In 2010, Bin Zhu and Tiande Shou reported that V4 has a positive correlation effect on the orientation selectivity of V1 (Shou, 2010b). Further, Jianbo Xiao and Xin Huang discovered the characteristics of MT cells for distinguishing complex orientations, indicating their great significance for the extraction of multiple movement directions (Xiao & Huang, 2015). Altogether, these experimental results have contributed considerably to understanding the basic principles of visual information processing (Li, 2014; Marr, 2010; Yin, Li, & Wu, 2018).

In parallel with neurobiological experiments, a number of neural computational models for the visual system have been likewise put forward. As early as in 1982, Marr firstly introduced a relatively comprehensive theory of visual computing informed by research grounded in neurobiology (Marr, 1982). He argued that visual cognition obtains “what” and “where” information through the “seeing” behavior and that the brain follows the hierarchical processing of visual information and the bottom-up principle. Such findings

are deemed to lay the groundwork for research in subsequent years. In 1999, Riesenhuber and Poggio proposed a model named "HMAX" based on V1 cells, mimicking the neural mapping from simple cells to complex cells in V1 (Riesenhuber & Poggio, 1999). In 2001, the widespread computing models of visual attention were brought to the fore, consisting of environmental stimuli saliency, saliency map, inhibition of return, attention and eye movements, scene understanding and object recognition, etc. These models enlarged the knowledge base concerning the neurobiological mechanisms of visual attention (Itti & Koch, 2001). In 2003, Li Zhaoping explored the segmentation and contour enhancement of V1 cells from the perspective of the computational model (Li, 2003). In 2006, Schölkopf and colleagues proposed a bottom-up model of visual saliency based on bottom-up attention, which was then employed calculate the visual saliency map in the corresponding scene (Schölkopf, Platt, & Hofmann, 2006). In 2011, Meng Xianglin and Wang Zhengzhi provided a model of enhancement strategy for region of interest based on attentional shroud, which possesses physiological and psychological rationality and can be used for region segmentation, target recognition, and scene analysis (Meng & Wang, 2011). In 2014, George et al. proposed a model for texture inhibition and contour enhancement based on the antagonistic and reverse inhibition properties of simple cells in V1 (Azzopardi, Rodríguez-Sánchez, Piater, & Petkov, 2014). In the same year, Jeroen et al. proposed a recurrent motion model (RMM) based on the response of the preferred orientations of MT cells, which can predict the perception of motion characteristics of MT cells (Joukes, Hartmann, & Krekelberg, 2014). In 2015, Chessa et al. proposed a V1-MT neural model for motion estimation, simulating the primary motion pathway of V1 to MT (Chessa, Sabatini, & Solari, 2016). More specifically, a two-dimensional Gabor filter was used to simulate the RFs of simple cells in V1, followed by obtaining the MT cells model through the weighted combination of V1 cells response and regularization, and subsequently applied to motion estimation. In 2017, Klaus et al. constructed an interference model of working memory about visual object feature information, based on four continuous-reproduction experimental data about working memory of color and direction (Oberauer & Lin, 2017). It is concluded that continuous visual information and discrete visual information have the same mechanisms of cue-based retrieval and interference. The findings thus paved the way for developing a unified theory of working memory in verbal, spatial, and visual information.

As the above reviews suggested, exploration into the visual information mechanisms have went through a long developmental period giving rise to a substantial amount of scientific achievements both in the field of neurobiological experiments and of computational neuroscience. Nevertheless, there has been a lack of a well-established theory to elucidate the significant phenomenon of visual information dynamic degradation in the visual nervous system.

Clearly, the visual information dynamic degradation occurs in the visual system. According to the experimental data provided by Anderson and Raichle (Anderson, Van Essen, & Olshausen, 2005; Raichle, 2010), the real world can actually emanate unlimited visual information. However, in the visual nervous system, only about 10^{10} bits/sec are deposited in the retina, which can be translated as nearly 1 million axons in each nerve from the neurobiological point of view. As a result of this limited number of axons in the optic nerves, only about 6×10^{6} bits/sec leave the retina, and only 10^{4} bits/sec can get to Layer IV of V1 (Raichle, 2010; Zhong & Wang, 2020). It can be seen that during the process of transmitting from the retina to Layer IV of V1, the visual information is reduced by about 10^{-6} times. Yet, the dynamic degradation cannot prevent visual cortex from gaining a complete visual perception of the real world.

Previous research shows that there is a convolution calculation approach for the pathway of retina-LGN-V1 (Zhong & Wang, 2020). Not only does it contain significant visual information dynamic degradation, it can also extract the edge features efficiently according to the principle of energy minimization of brain activity. Moreover, the computational model proposed in accordance with such findings provides quantitative methods to understand the neural mechanisms of the dynamic degradation mapping from the retina to V1, which can produce results that match the experimental data noted above.

As we mentioned earlier, however, the mechanism of visual information mapping from V1 to V2 still remains unclear as regards, the existence (or not) of degradation during the mapping process and the way in which such visual information can be quantitatively analyzed (Semedo, Zandvakili, Machens, Byron, & Kohn, 2019; Zhaoping, 2019). These are vital to understanding visual information processing in higher-order cortices.

Due to a lack of available models to address these questions, we established a computational model in the current paper to quantitatively predict and analyze the visual information dynamic degradation based on the mapping from V1 to V2. The study was informed by the convolution calculation approach for the pathway of retina-LGN-V1 (Zhong & Wang, 2020), the theory of convolutional neural networks (CNN) (LeCun, Bengio, & Hinton, 2015) and anatomical architecture between V1 and V2 (Gazzaniga, Ivry, & Mangun, 2019).

The novelty of this study mainly consists in the following three respects.

First, CNN is directly inspired by the classic notions of simple cells and complex cells in the visual system, and the overall architecture relies on the LGN-V1-V2 hierarchy in the visual cortex (LeCun et al., 2015). Drawing the lesson from CNN, we have built a computational model in the previous study, the results of which were consistent with experimental data and proved its feasibility. Therefore, we extend that model based on the anatomical architecture between V1 and V2, which is of great value to research on visual information processing from a new theoretical perspective.

Second, the computational model proposed by the current paper, which includes 6 layers simulating the levels of Photoreceptors, Ganglion cells, LGN, V1, and V2, mimics the visual information processing. The results indicate there still exists convolution calculation and a slight degree of dynamic degradation in V1-V2. Specifically, the visual information of V2 is 0.18 times that of V1, which offers us a precise understanding of the visual information mapping mechanism from V1 to V2. In addition, the computational results will make up for the lack of experimental data of V1-V2.

Lastly, the results demonstrate that although the RFs of V2 have strong responses to the “corner” of the visual image (Hosoya & Hyvärinen, 2015), they do not extract the feature information to any further degree. Therefore, it can be concluded that the significant dynamic degradation occurs in the pathway of the retina to V1. In other words, the novel visual information from the real world is entirely processed in the early visual areas and primarily processed in retina-LGN-V1. On the other hand, following the principle of synaptic plasticity, the RFs of V2 can accurately respond to and encode the scarce “corner” information about the real world. The contour detection (edge and corner detection) of visual perception in natural scenes only uses lower-order areas’ visual information.

2 Methods

2.1 The Visual Information Changes from Retina to V1

The visual system grants animals the capability to perceive the real world (Gazzaniga et al., 2019). In the ventral pathway of the visual cortex, the form perception is gradually improved with respect to the cortical hierarchy of low-order to high-order (Hatori, Mashita, & Sakai, 2016). In V1, V2, and V4, their RFs are selective for orientations, angles, and curvatures, respectively.

The light reaching the retina, and then mapping to LGN, and V1, the sequence of visual information processing follows **Figure 1**, as shown below (Zhong & Wang, 2020).

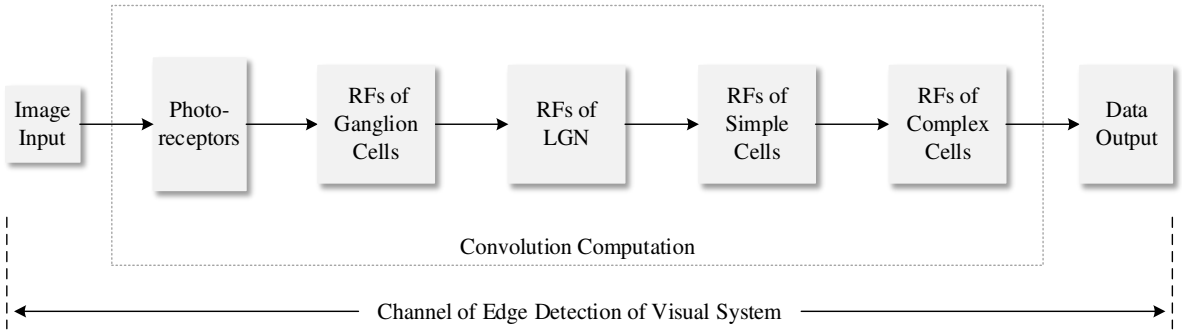

Fig.1 *The visual information processing from retina to V1.*

The pathway of the retina to V1 is a one-to-one neural mapping (Zhao et al., 2010). The photoreceptor converts the external light signals into bioelectrical signals and delivers them to ganglion cells, which finally are transmitted to V1 through LGN. In this system, about 10^{10} bits/sec are deposited in the retina; only 10^4 bits/sec can get to V1. Obviously, the visual information changes from the retina to V1 is dynamic degradation.

The type model of ganglion cell is On-center, which is shown in **Figure 2**.

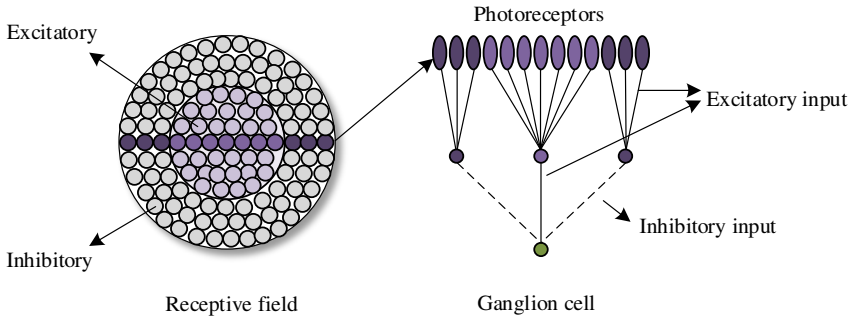

Fig.2 *On-center model of ganglion cell.*

If the stimuli of light are located at the inner circle of RFs of photoreceptors, the ganglion cells generate higher frequency action potentials. If the stimuli are located at the outer circle of RFs, the ganglion cells inhibit frequency action potentials. Consequently, the central stimulation response and the peripheral stimulation response are mutually offset, leading to the discovery that the ganglion cells are very sensitive to

the difference in brightness in RFs (Obara, O'Hashi, & Tanifuji, 2017). For this reason, the visual information dynamic degradation occurs (Raichle, 2010; Zhong & Wang, 2020).

The architectures of RFs of LGN are the same as those of RFs of ganglion cells, which include two concentric circles (Gazzaniga et al., 2019). After the processing of LGN, the visual information is transmitted to V1. Similarly, LGN also can identify corresponding features. These characteristics make the visual information further degrade after the processing of LGN.

Both the simple cells and complex cells in V1 display a strong response to the specific preferred orientation. The architecture of simple cells is very similar to that of the Gabor filter (Ringach, 2002), as shown in **Figure 3**. Complex cells have no requirement for specific locations and are the abstraction of simple cells. These characteristics of V1 cells further strengthen the capability of feature detection, which also degrades the visual information.

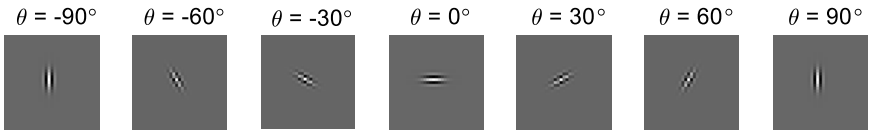

Fig.3 *Simple cells for different preferred orientations.*

2.2 The Visual Information Changes from V1 to V2

The previous section 2.1 briefly introduced and analyzed the reason for visual information dynamic degradation from the retina to V1 in visual nervous system. Nevertheless, the visual information remains unknown for the changes in transmission from V1 to V2, and for the changes in ventral pathway transmission as the cortical order increases.

Sparse coding theory is a critical approach in visual information processing. Due to the restriction of energy metabolism during brain information processing and signal transmission, the number of neurons that process large amounts of visual information is very few (Hosoya & Hyvärinen, 2015). To some extent, the activity of simple cells in V1 can be summed as a linear function of RFs in a small spatial position. We could utilize the Gabor function to represent the characteristics of the two-dimensional mapping of simple cells (Olshausen & Field, 1996). The complex cells are regarded as the abstraction of simple cells. There is no significant difference directly from the morphological perceptive of simple cells and complex cells, which seem to be the same type of cells (Goodfellow, Bengio, & Courville, 2016). Some research results have shown that the functional classification of simple cells and complex cells is not static, and their functions can mutually transform into each other (Shou, 2010a). From the perspective of the computational model, tuning parameters achieve continuous behavior from simple cells to complex cells.

V2 cells have a characteristic of selectivity for the corners (Banich & Compton, 2018), comprising two different lines from end to end, each direction of which is derived from V1 cells. Consequently, V2 cells can be represented as two weighed Gabor filters (Zhao et al., 2010; Ziemba, Freeman, Simoncelli, & Movshon, 2018).

According to the hierarchical hypothesis model of the primary visual cortex proposed by Hubel and

Wiesel (Hubel & Wiesel, 1962), the external information arriving at the visual system abides by the principles of the pathway of retina-LGN-V1-V2. Concerning the information separation and processing model proposed by Livingstone and Hubel (Livingstone & Hubel, 1987), the shape, color, motion, and stereopsis are separated in V1 and V2 during the information processing of retina-LGN-V1-V2. Since we focus on the visual information changes, we pay greater attention to the shape. Hosoya and Hyvärinen have proposed a model based on a 3-layer network consisting of simple cells, complex cells, and V2 cells (Hosoya & Hyvärinen, 2015). Accordingly, we contend that the visual information processing of retina-LGN-V1-V2 in the visual system, shown in **Figure 4**, can be represented by a structural schematical diagram, as shown in **Figure 5**.

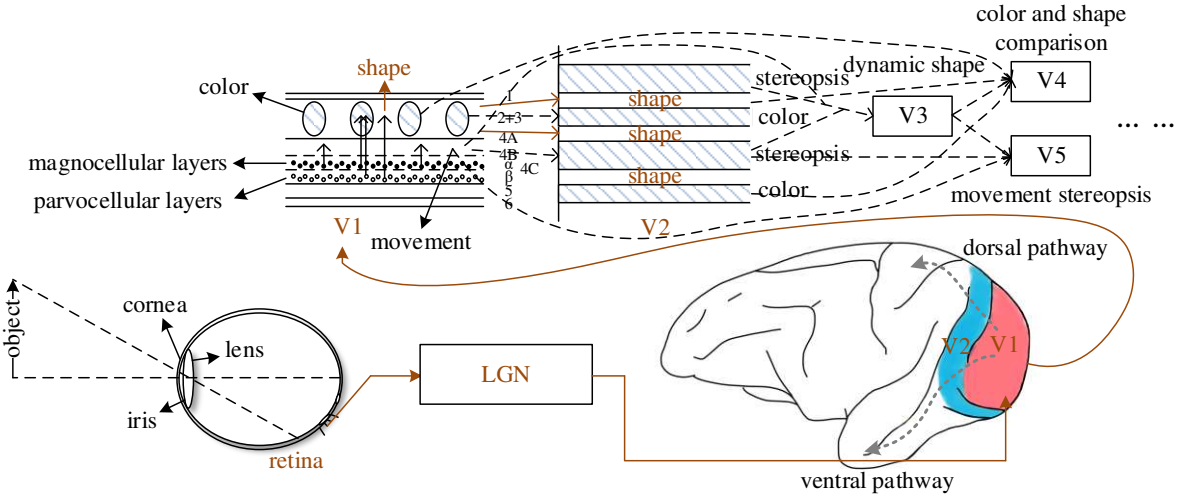

Fig.4 The diagram of visual information processing and transmission.

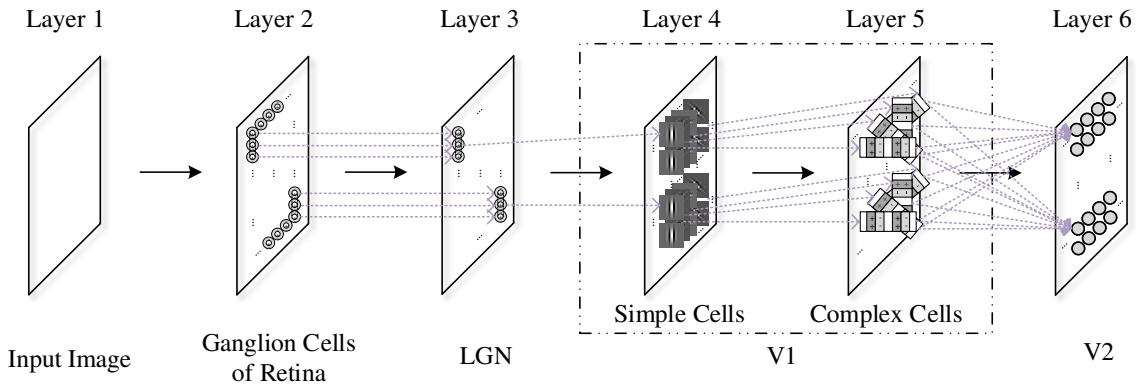

***Fig.5** The structural diagram of visual information from retina to V2.*

Figure 4 shows that the red-brown line represents the external information transmission by neural mapping from the retina to V2. With the model designed on the basis of **Figure 4** and neurobiological experiments (Hosoya & Hyvärinen, 2015; Lee, Ekanadham, & Ng, 2008), we have established a structural schematic diagram of visual information transmitting to V2, as shown in **Figure 5**. **Figure 5** allows for the calculation of the visual information changes from V1 to V2.

2.3 The Analysis of Visual Information Changes from V1 to V2

The visual information changes from V1 to V2 has long puzzled neuroscientists. In other words, there is no available method to quantitatively analyze the visual information changes from V1 to V2 from the perspective of neurobiological experiments or computational models, which hinders a fuller grasp of the mechanisms of visual information processing. Literature suggests that the edge detection channel of the visual system, which is the functional channel of retina-LGN-V1, has the characteristic of one-to-one neural mapping (Zhong & Wang, 2020). The mapping mechanism from the retina to V1 is closely related to the convolution calculation, which partly causes significant dynamic degradation. The EDMRV1 model is established based on the pathway of photoreceptor-ganglion cell-LGN-V1. The simulation results turned out to fit well with the experimental data provided by Anderson (Anderson et al., 2005), and clearly explained the dynamic degradation phenomenon, as shown in **Figure 6**.

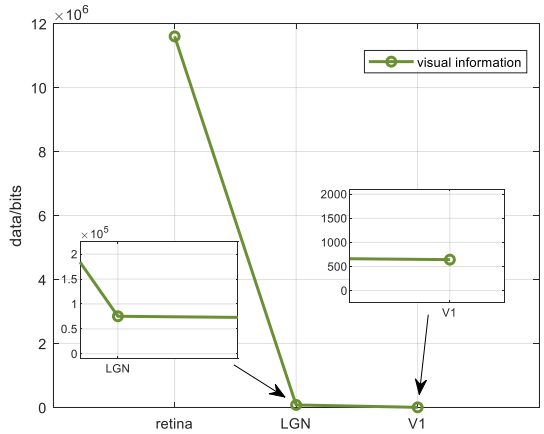

Fig.6 *The visual information dynamic degradation of photoreceptor-ganglion cell-LGN-V1 based on EDMRV1 model (Zhong & Wang, 2020).* The line that represents the visual information of LGN and V1 is very close to the x-axis. For more details, we have zoomed in to clarify.

Since the visual information processing is processed by the RFs of simple cells and complex cells in V1, the processed information is directly output to V2 cells. According to the neural mapping from V1 to V2, it can be argued that one RF of V2 is weighted by two RFs of V1, which are the same or different preferred orientations (Hosoya & Hyvärinen, 2015). Hence, the visual information is in fact transmitted in the way of connection. Given this connection, we established a visual information detection model based on V2, which could predict and calculate the visual information changes in V2 through a quantitative analysis.

As existing research uncovers (Zhong & Wang, 2020), in the pathway of photoreceptor-ganglion cell-LGN-V1, the visual information hierarchical transmission from low-order to high-order visual cortex follows the convolution calculation. It is also the main reason for the visual information changes from the retina to LGN to V1. The RFs of V2 are constructed by the combination of RFs of V1 (Minami & Naokazu, 2011), which accounts for the existence of convolution calculation exists in the neural mapping from V1 to V2. Therefore, it is reasonable and feasible to use the photoreceptor-ganglion cell-LGN-V1-V2 model to predict and calculate the visual information changes.

Due to the intricate connections between neurons, the characteristic of connections in different RFs is closely related to spike timing-dependent plasticity (STDP) (Beyeler, Dutt, & Krichmar, 2013; Kim & Lim, 2019), which is also known as pulse-time-dependent plasticity. The connection characteristics are also tightly linked with the orientation selectivity of RFs (Carver, Roth, Cowan, & Fortune, 2008). STDP comprises two types: long-term potentiation (LTP) and long-term depression (LTD) (Gazzaniga et al., 2019). The relationship between the sequence of firing and the connection strength determines the detection of image feature information by RFs.

1) In the area of edges of the image, the presynaptic and postsynaptic neurons produce synchronous and positive discharge with high-probability under the LTP effect. At this time, the synaptic connection is continuously strengthened, as expressed in the following:

$$\text{potentiation}(i,\ j) = t_{\text{post}}(i,\ j) \times \left(1 + \text{synapse}(i,\ j) \times e^{-\left|t_{\text{pre}}(i,\ j) - t_{\text{post}}(i,\ j)\right|}\right),\ t_{\text{pre}} < t_{\text{post}}, \tag{1}$$

where potentiation(i, j) represents the decoding information of the image features after the LTP effect in STDP, synapse(i, j) indicates the strength of the synapse connection.

2) In the none-edge area, the presynaptic and postsynaptic neurons produce non-synchronous and high-probability non-positive discharges under the LTD effect. At this point, the synaptic connections are constantly suppressed, expressed by the following:

$$\text{depression}(i,\ j) = t_{\text{post}}(i,\ j) \times \left(1 - \text{synapse}(i,\ j) \times e^{-\left|t_{\text{pre}}(i,\ j) - t_{\text{post}}(i,\ j)\right|}\right),\ else, \tag{2}$$

where depression(i, j) represents image edge decoding information after the LTD effect in STDP. Concerning Equation (1) and (2), $t_{\text{pre}}<t_{\text{post}}$ indicates that neurons pre- and postsynaptic neurons are positively discharged. Reversely, $t_{\text{pre}}\geq t_{\text{post}}$ indicates that neurons pre- and postsynaptic neurons are non-positively discharged. The results of STDP rule as shown in the following **Figure 7**,

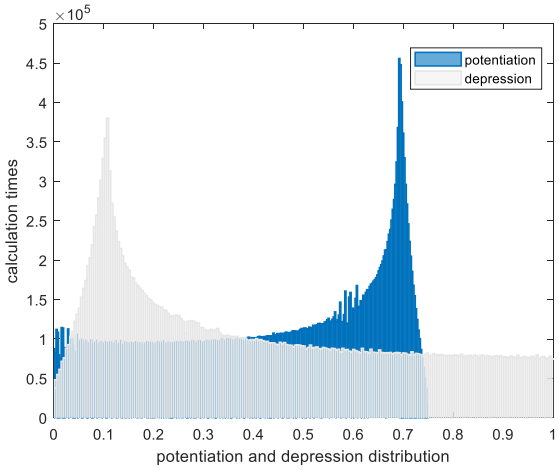

Fig. 7 *STDP rule.* The potentiation and depression distribution of weights. The x-axis indicates the results of potentiation(i, j) and depression(i, j); the y-axis, the calculation times, which means the higher value of the y-axis indicates that the higher frequency of appearance of the corresponding x value.

To this end, considering **Figure 5**, in order to advance research on the changes of visual information from V1 to V2 and the mechanism of visual information processing of V2, we designed a 6-layer feedforward network model, which is a predictive model for visual information changes in V2 (PMVICV2).

In our proposed model, Layer 1 represents the photoreceptor of the retina. The real world's information is transmitted to the photoreceptor after being refracted by the lens and then converted into a bioelectric signal. At this point, the size of the entire image on the retina is denoted as A, which depends on the specific experiment subjects. It is assumed that A could be divided into $M \times N$ patches. We define the image information on the photoreceptor as $I(i, j)$ (i=1, 2, 3, ..., M; j=1, 2, 3, ..., N), then:

$$A = \sum_{j=1}^{N} \sum_{i=1}^{M} I_{i,j}(a), \tag{3}$$

$$a = \Delta i \times \Delta j . \tag{4}$$

Layer 2 represents the RFs of ganglion cells; we define the visual information in ganglion cells as $I_2(i, j)$, each RF of ganglion cells defined as a. Being processed by horizontal cells and bipolar cells, $I_2(i, j)$ is the visual information transmitted by $I_1(i, j)$ to the ganglion cells. One ganglion cell will obtain signal inputs of 10^3~10^4 photoreceptors (Zhao et al., 2010). Suppose that the RF of a ganglion cell, shown in **Figure 2**, can be represented by DOG (i, j), the antagonism of the outer-circle and inner-circle can be described by the difference of two Gaussians (DOG) (Shou, 2010a)

$$DOG(i, j)_{\text{ganglion cell}} = DOG_1(i, j) - DOG_2(i, j) = k_c e^{\left(-\frac{i^2+j^2}{2r_c^2}\right)} - k_s e^{\left(-\frac{i^2+j^2}{2r_s^2}\right)}, \tag{5}$$

where DOG_1 represents the inner-circle, and DOG_2 represents the outer-circle. K_c and K_s respectively indicate the maximum sensitivity of the central and peripheral areas of the RF. r_c and r_s represent the radius concerning the maximum sensitivity of central and peripheral areas of the RF when they drop to e^{-1}.

Each $I_{i,j}(a)$ contains the visual information of the corresponding RF and feature information of some images. In other words, the RF of each ganglion cell has the corresponding $I_{i,j}(a)$. Such neural mapping exists extensively in the visual system. A patch $I_{i,j}(a)$ activates the corresponding ganglion cell and triggers its higher frequency action potential.

At this point, the visual information on the RF of the ganglion cell is $I_2(i, j)$, shown in the following:

$$I_2(i, j) = (I_1 * DOG_{\text{ganglion cell}})(i, j). \tag{6}$$

In Layer 3, the ganglion cells in Layer 2 have processed the visual information, which is transmitted to LGN. The RF of LGN is divided into two antagonistic areas, of which the structure and function are very similar to that of the ganglion cell (Gazzaniga et al., 2019). To this end, we still use the *DOG* model for representation. Then, suppose the visual information on the RF of LGN in Layer 3 is $I_3(i, j)$, which is mapped from Layer 2, shown in the following:

$$I_3(i, j) = (I_2 * DOG_{\text{LGN}})(i, j). \tag{7}$$

In Layer 4, the simple cells in V1 have orientation selectivity for the features on the image (Liu et al., 2010); that is, they have strong selectivity for features with specific directions, at which point the corresponding neuron responds strongest in the direction, shown as a two-dimensional Gabor function (Shou, 2010a):

$$G_{\lambda,\theta,\psi,\sigma,\gamma}(i,j) = e^{-\frac{i'^2+\gamma^2 j'^2}{2\sigma^2}} \cos\left(2\pi \frac{i'}{\lambda} + \psi\right), \tag{8}$$

$$\begin{cases} i' = i\cos\theta + j\sin\theta \\ j' = -i\sin\theta + j\cos\theta \end{cases}. \tag{9}$$

Equation (8) is the product of a Gaussian function and a cosine function. λ is the wavelength, which directly affects the filter scale of the filter. θ is the direction of the filter. ψ is the phase shift of the tuning function. γ is the ratio of spatial vertical to horizontal. σ is the variance of the Gaussian filter.

The RFs in V1 prefer different orientations (Gazzaniga et al., 2019); we sampled every 60° with 3 orientations of RFs, considering that the V1 area is not the emphasis of the current research. Then, we supposed the visual information on the RFs of the simple cells in Layer 4 is $I_4(i, j)$, which came after the neural mapping of Layer 3, shown in the following:

$$I_4(i,j) = (I_3 * Gabor_{v1})(i,j). \tag{10}$$

In Layer 5, complex cells originate from the inputs of simple cells at the same orientation but at different locations, which means the abstraction of RFs of simple cells. Since the RFs of complex cells have no clear antagonist area, there is no strict requirement for the location as the orientation selection. Simple cells and complex cells can sometimes be converted functionally. The image information on Layer 5 can be written as $I_5(i, j)$. Therefore, $I_5(i, j)$ and $I_4(i, j)$ are considered the same function.

In Layer 6, the visual information on the RFs in V2 is the neural mapping from Layer 5, recorded as $I_6(i, j)$. The RF in V2 is composed of two RFs in V1. Each RF's preferred orientation could be the same or different. The RF in V2 is selective for the angle profile, shown in the following Equation:

$$I_6(i,j) = \left(I_5 * \left(Gabor_{v1_1} + Gabor_{v1_2}\right)\right)(i,j), \tag{11}$$

where the visual information reaches Layer 5, that is, the neural mapping from V1 to V2, $I_5(i, j)$ is equivalent that a series of stimuli react to the different RFs in V2. Therefore, the RFs in V2 extract the corresponding features according to the different strengths of the stimuli. To this end, the model of Layer 6 is composed of two RFs in V1 with the preferred direction. Such a combination forms an angle, the value range of which is [0, 360] in degree. Also, each angle has a direction, the value range of which is also [0, 360]. The unit is degree.

The RFs in V2 have different preferred angles according to varying degrees and directions. In the current study, each 30° can be used as the sample angle. As such, 12 different angles and directions are shown in

Figure 8.

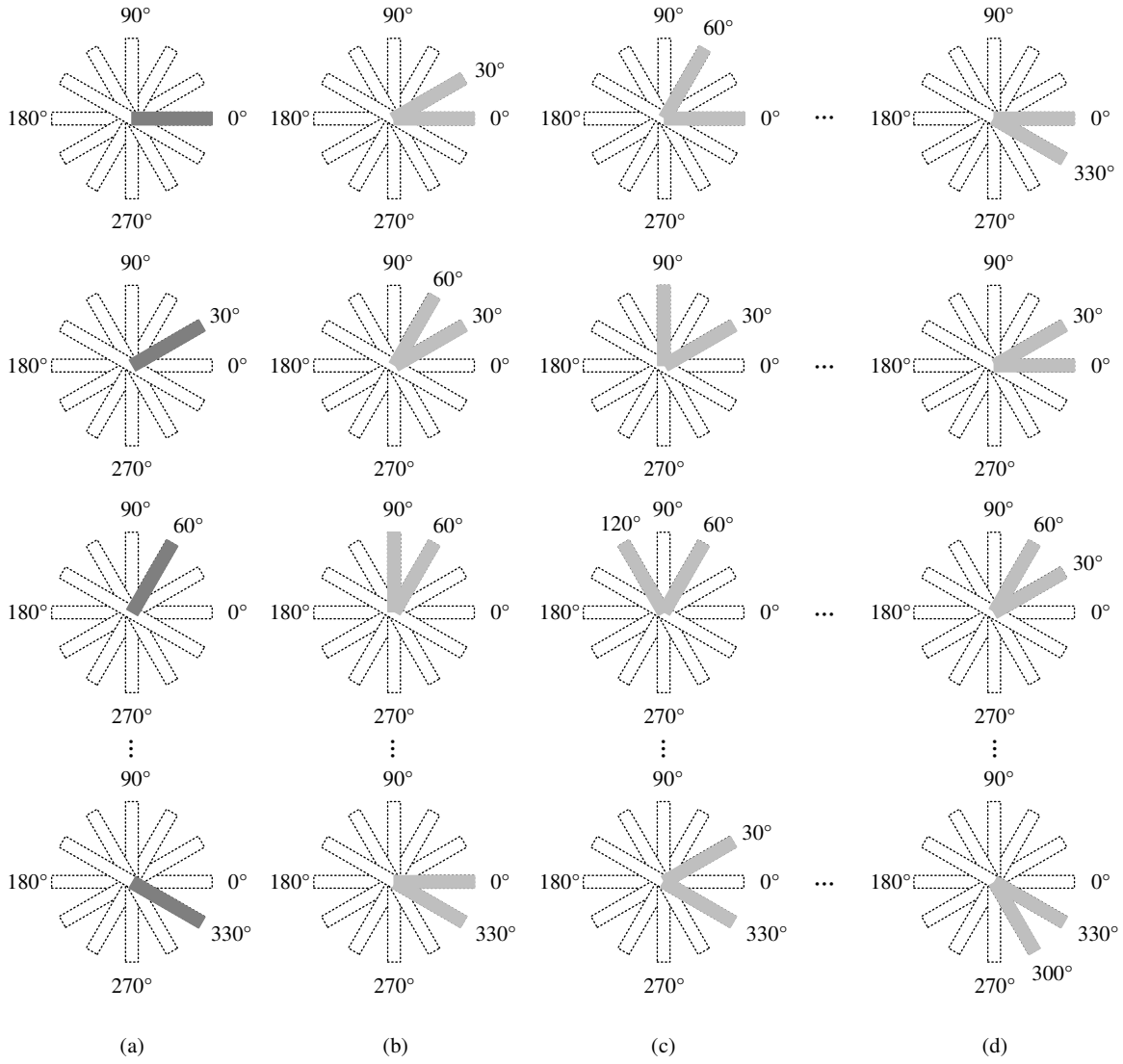

Fig.8 *The RFs (two grey sides) in V2 with varying angles and directions.* **a.** The angles of RFs equal 0° in different directions. The dark grey indicates two sides of RFs are overlapped. **b.** The angles of RFs equal 30° in different directions. **c.** The angles of RFs equal 60° in different directions. **d.** The angles of RFs equal 330° in different directions.

According to the first row of **Figure 8**, each angle has two sides; one of them is fixed, the other is rotated. These two form the shape with different degrees. In the first column, the shape of the angle is fixed; the rotation forms different directions of angles. Each side of the angle is an RF in V1, which has a specific orientation preference. Among the RFs in V2, angles in the first column are 0, as shown in **(a) of Figure 8**. The second column of RFs has angles with 30°, as shown in **(b) of Figure 8**. Angles in the third column are 60°, as shown in **(c) of Figure 8**. The fourth has angles with 330°, as shown in **(d) of Figure 8**.

The angles with different degrees and directions in Layer 6 are defined as follows:

1) $angle_{\text{size}}$ is indicated as the following Equation (12):

$$\left\{angle_{\text{size}} \mid angle_{\text{size}} = 30° \times n, n \in [0, 11] \text{and } n \in N\right\}. \tag{12}$$

2) $angle_{\text{orientation}}$ is shown as the following Equation (13):

$$\{angle_{\text{orientation}} \mid angle_{\text{orientation}} = (2n+1)\times 15°, n \in [0,11] \text{ and } n \in N\}. \quad (13)$$

On the effect of visual information stimuli in Layer 5, Layer 6 performs convolution calculation with RFs at varying angles and directions in V2, and finally obtains RFs responses in V2.

3 Results and Analyses

3.1 Simulation

According to the description of the above PMVICV2 model, the following four diverse scenarios are used as experimental examples. In the V1 area, the sampling angle is 60°; in the V2 area, the sampling angle is 30°, 60°, 90°, 120°, 150°, 180°, sequentially. The phenomenon of dynamic degradation exists on the pathway of photoreceptor-ganglion cell-LGN-V1 and the pathway of V1-V2; ~~the experimental results are challenging to identify with the human eye. Accordingly, the brightness and contrast value of the experimental images were respectively reduced and increased by 40%.~~ Each pixel of images was encoded in one byte.

3.1.1 Experiment of the portrayal of Lena

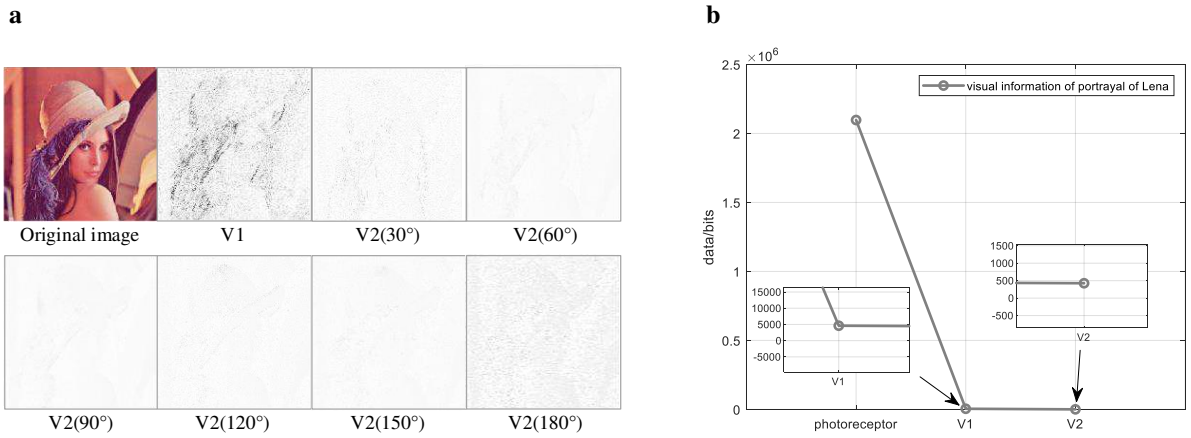

Fig. 9 *Lina image and PMVICV2 model responses.* **a.** Lina original image and the processed image in V1(60°) and V2 (30°, 60°, 90°, 120°, 150°, 180° with the same direction). **b.** The results of (a): dynamic degradation of Lina image processed by PMVICV2 model. The line that represents the visual information of V1 and V2 is very close to the x-axis. For more details, we have zoomed in to clarify.

As shown in the **(b) of Figure 9**, the picture of Lena (original image), of which resolution was 512×512, was utilized as the experimental object. The optical signal reached the photoreceptors, the visual information of which was 2.10×10^6 bits. Subsequently, the visual information was processed by cones and rods and then reached ganglion cells for processing. Since the RFs in ganglion cells had antagonistic properties that are

highly sensitive to the change of light and dark, the edge feature information of the image could be detected. The dynamic degradation occurred after the visual information was transmitted to LGN for processing. Simple cells with different preferred orientations are defined as $\theta_{60°,120°,180°}$, which actively responded to the image information and recognized the edge feature information in the specific orientation. The visual information in V1 was 1.34×10^3 bits，1.07×10^3 bits and 1.14×10^4 bits, respectively, which was about 6.39×10^{-4} times, 5.11×10^{-4} times, and 5.43×10^{-3} times that of the photoreceptors. The visual information in V1 was obtained by the processing of RFs of ganglion cells and those of LGN. Finally, the visual information was transmitted from V1 to V2. The RFs in V2 have a strong response to the different corresponding angles and directions, which can identify the image feature information. These angles are denoted as $angle_{\text{size}}$, shown in Equation (14):

$$\left\{angle_{\text{size}} \mid angle_{\text{size}} = 30°\times n, n \in [1, 6] \text{ and } n \in N\right\}. \tag{14}$$

The image after processing is shown in **(a) of Figure 9**. The visual information in V2 is shown in Table 1:

Table 1 Visual information in V2 of the experiment of Lina (Unit: bits)

Orientation\Size	30°	60°	90°	120°	150°	180°
1	6.33×10^2	92.5	11.85	29.75	2.66×10^2	1.23×10^3
2	2.66×10^2	29.75	11.85	92.5	6.33×10^2	1.81×10^3
3	6.33×10^2	92.5	11.85	29.75	2.66×10^2	1.23×10^3
4	2.66×10^2	29.75	11.85	92.5	6.33×10^2	1.81×10^3
Average	4.50×10^2	61.13	11.85	61.13	4.50×10^2	1.52×10^3

The comparison of visual information between photoreceptors and V2, shown as the following:

Table 2 Relationship between photoreceptors and V2 of the experiment of Lina

Orientation\Degraded rate\Size	30°	60°	90°	120°	150°	180°
1	3.02×10^{-4}	4.41×10^{-5}	5.65×10^{-6}	1.42×10^{-5}	1.27×10^{-4}	5.89×10^{-4}
2	1.27×10^{-4}	1.42×10^{-5}	5.65×10^{-6}	4.41×10^{-5}	3.02×10^{-4}	8.64×10^{-4}
3	3.02×10^{-4}	4.41×10^{-5}	5.65×10^{-6}	1.42×10^{-5}	1.27×10^{-4}	5.89×10^{-4}
4	1.27×10^{-4}	1.42×10^{-5}	5.65×10^{-6}	4.41×10^{-5}	3.02×10^{-4}	8.64×10^{-4}
Average	2.15×10^{-4}	2.92×10^{-5}	5.65×10^{-6}	2.92×10^{-5}	2.15×10^{-4}	7.27×10^{-4}

From the above analysis, considering the image of Lena as the experimental object, we have indicated that the changes of visual information from the retina to V1 and V2, shown in **(b) of Figure 9**. It can be recognized that the average value of visual information of photoreceptors was 2.10×10^6 bits; the average value of V1 was 4.60×10^3 bits; the average value of V2 was 4.26×10^2 bits. These values demonstrated that the visual information degrades significantly from photoreceptors to V1. The visual information of V1 was 2.20×10^{-3} times that of the photoreceptor. Nevertheless, during the processing from V1 to V2, the dynamic degradation already existed but was scanty; the visual information of V2 was 9.25×10^{-2} times that of V1.

3.1.2 Experiment of the island of Manhattan

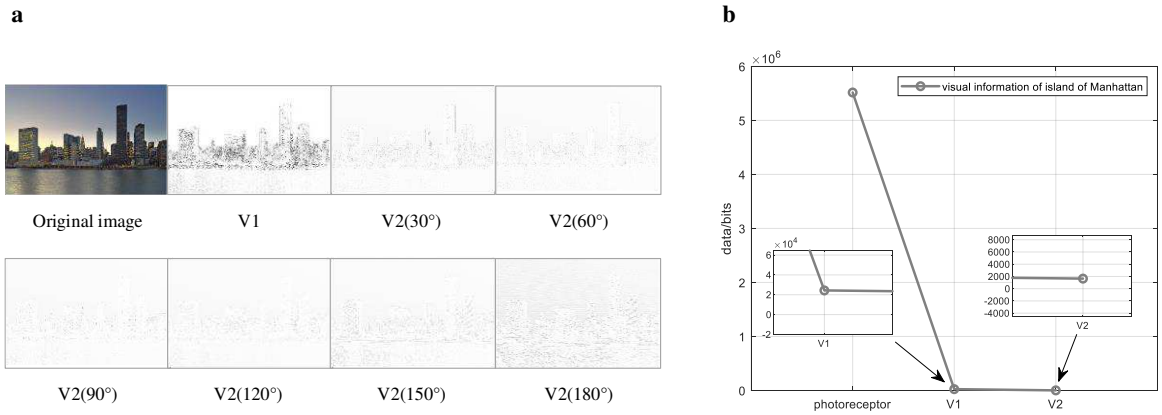

Fig. 10 *Manhattan image and PMVICV2 model responses.* **a.** Manhattan original image and the processed image in V1(60°) and V2 (30°, 60°, 90°, 120°, 150°, 180° with the same direction). **b.** The results of (a): dynamic degradation of Manhattan image processed by PMVICV2 model. The line that represents the visual information of V1 and V2 is very close to the x-axis. For more details, we have zoomed in to clarify.

As shown in **(a) of Figure 10**, the Manhattan image (original image), of which resolution was 1023×674, was used as the experimental object. The photoreceptor received the optical signal, the visual information of which was 5.52×10^6 bits. Afterward, the visual information was processed by cones and rods, and then passed to ganglion cells for processing. As the RFs in ganglion cells had antagonistic properties, the edge features of the image could be detected. The dynamic degradation occurred after the visual information was transmitted to LGN for processing. Simple cells with different preferred orientations are defined as $\theta_{60°,120°,180°}$, which actively responded to the image information and recognized the edge feature information in the specific orientation. The visual information in V1 was 3.52×10^3 bits，3.36×10^3 bits and 6.62×10^4 bits, respectively, which was about 6.39×10^{-4} times, 6.09×10^{-4} times, and 1.20×10^{-2} times that of the photoreceptors. The visual information in V1 was obtained by the processing of RFs of ganglion cells and those of LGN. Finally, the visual information was transmitted from V1 to V2. The RFs in V2 had a strong response to the different corresponding angles and directions denoted as $angle_{\text{size}}$, which can identify the image feature information, shown in Equation (14). The processed image is shown in **(a) of Figure 10**. Lastly, the visual information in V2 is shown in Table 3:

Table 3 Visual information in V2 of the experiment of Manhattan (Unit: bits)

Orientation\Size	30°	60°	90°	120°	150°	180°
1	1.99×10^3	5.07×10^2	1.58×10^2	4.94×10^2	2.65×10^3	4.72×10^3
2	2.65×10^3	4.94×10^2	1.58×10^2	5.07×10^2	1.99×10^3	3.46×10^3
3	1.99×10^3	5.07×10^2	1.58×10^2	4.94×10^2	2.65×10^3	4.72×10^3
4	2.65×10^3	4.94×10^2	1.58×10^2	5.07×10^2	1.99×10^3	3.46×10^3
Average	2.32×10^3	5.01×10^2	1.58×10^2	5.01×10^2	2.32×10^3	4.09×10^3

The comparison of visual information between photoreceptors and V2 is shown in the following:

Table 4 Relationship between photoreceptors and V2 of the experiment of Manhattan

Orientation\Degraded rate\Size	30°	60°	90°	120°	150°	180°
1	3.61×10^{-4}	9.20×10^{-5}	2.87×10^{-5}	8.95×10^{-5}	4.80×10^{-4}	8.56×10^{-4}
2	4.80×10^{-4}	8.95×10^{-5}	2.87×10^{-5}	9.20×10^{-5}	3.61×10^{-4}	6.27×10^{-4}
3	3.61×10^{-4}	9.20×10^{-5}	2.87×10^{-5}	8.95×10^{-5}	4.80×10^{-4}	8.56×10^{-4}
4	4.80×10^{-4}	8.95×10^{-5}	2.87×10^{-5}	9.20×10^{-5}	3.61×10^{-4}	6.27×10^{-4}
Average	4.20×10^{-4}	9.07×10^{-5}	2.87×10^{-5}	9.07×10^{-5}	4.20×10^{-4}	7.41×10^{-4}

From the above analysis, taking the image of the island of Manhattan as the experimental object, we have indicated that the visual information changes from the retina to V1 and V2, as shown in **(b) of Figure 10**. It can be recognized that the average value of visual information of photoreceptors was 5.52×10^{6} bits; the average value of V1 was 2.43×10^{4} bits; the average value of V2 was 1.65×10^{3} bits. These values demonstrated that the visual information degraded significantly from photoreceptors to V1. The visual information in V1 was 4.41×10^{-3} times that in the photoreceptor. Nevertheless, during the processing from V1 to V2, the dynamic degradation already existed but was scanty; the visual information in V2 was 6.77×10^{-2} times that in V1.

3.1.3 Experiment of the harbor of Sydney

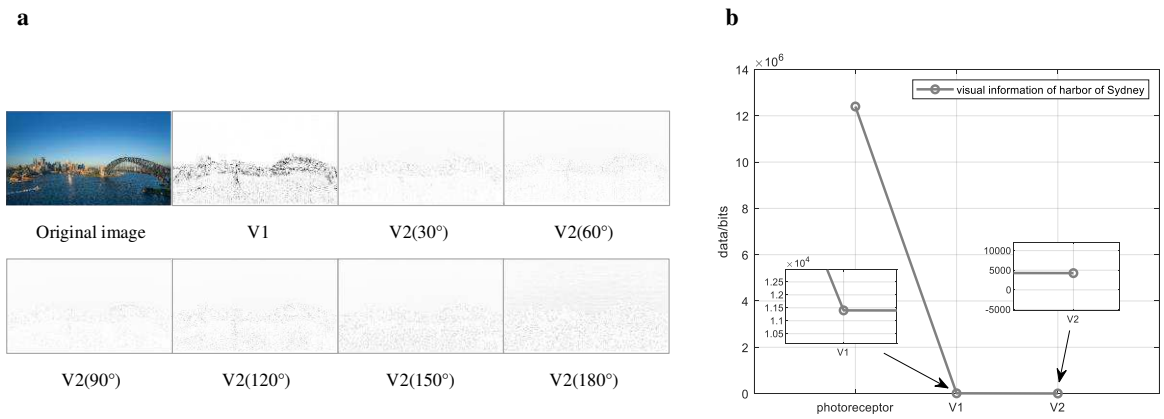

Fig. 11 *Sydney image and PMVICV2 model responses.* **a.** Sydney original image and the processed image in V1(60°) and V2 (30°, 60°, 90°, 120°, 150°, 180° with the same direction). **b.** The results of (a): dynamic degradation of Sydney image processed by PMVICV2 model. The line that represents the visual information of V1 and V2 is very close to the x-axis. For more details, we have zoomed in to clarify.

As shown in **(a) of Figure 11**, the Sydney image (original image), the resolution of which is 1663×934, was used as the experimental object. The photoreceptor received the optical signal of the visual information, which was 1.24×10^{7} bits. Subsequently, the visual information was processed by cones and rods, and then passed to ganglion cells. The RFs in ganglion cells easily detected the edge features of the image. The dynamic degradation occurred after the visual information was transmitted to LGN. Simple cells with different preferred orientations are defined as $\theta_{60°,120°,180°}$, the visual information of which in V1 was 7.17×

10^3 bits，6.83×10^3 bits and 2.02×10^4 bits, respectively, which was about 5.77×10^{-4} times, 5.50×10^{-4} times and 1.63×10^{-3} times that of the photoreceptors. The visual information in V1 was obtained by the processing of RFs of ganglion cells and LGN. Finally, the visual information was transmitted from V1 to V2. The RFs in V2 had a strong response to the different corresponding angles and directions denoted as $angle_{size}$, which can identify the image feature information, as shown in Equation (14). The processed image is shown in **(a) of Figure 11**. Lastly, the visual information in V2 is shown in Table 5:

Table 5 Visual information in V2 of the experiment of Sydney (Unit: bits)

Orientation\Size	30°	60°	90°	120°	150°	180°
1	3.24×10^3	9.33×10^2	3.48×10^2	2.60×10^3	9.43×10^3	1.26×10^4
2	9.43×10^3	2.60×10^3	3.48×10^2	9.33×10^2	3.24×10^3	5.37×10^3
3	3.24×10^3	9.33×10^2	3.48×10^2	2.60×10^3	9.43×10^3	1.26×10^4
4	9.43×10^3	2.60×10^3	3.48×10^2	9.33×10^2	3.24×10^3	5.37×10^3
Average	6.34×10^3	1.76×10^3	3.48×10^2	1.76×10^3	6.34×10^3	9.00×10^3

The comparison of visual information between photoreceptors and V2 is shown in the following:

Table 6 Relationship between photoreceptors and V2 of the experiment of Sydney

Orientation\Degraded rate\Size	30°	60°	90°	120°	150°	180°
1	2.61×10^{-4}	7.51×10^{-5}	2.80×10^{-5}	2.09×10^{-4}	7.59×10^{-4}	1.02×10^{-3}
2	7.59×10^{-4}	2.09×10^{-4}	2.80×10^{-5}	7.51×10^{-5}	2.61×10^{-4}	4.33×10^{-4}
3	2.61×10^{-4}	7.51×10^{-5}	2.80×10^{-5}	2.09×10^{-4}	7.59×10^{-4}	1.02×10^{-3}
4	7.59×10^{-4}	2.09×10^{-4}	2.80×10^{-5}	7.51×10^{-5}	2.61×10^{-4}	4.33×10^{-4}
Average	5.10×10^{-4}	1.42×10^{-4}	2.80×10^{-5}	1.42×10^{-4}	5.10×10^{-4}	7.24×10^{-4}

From the above analysis, taking the image of Sydney as the experimental object, we have shown that visual information changes from the retina to V1 and V2, as illustrated in **(b) of Figure 11**. It can be recognized that the average value of visual information of photoreceptors was 1.24×10^7 bits; the average value of V1 was 1.14×10^4 bits; the average value of V2 was 4.26×10^3 bits. These values demonstrated that the visual information degrades significantly from photoreceptors to V1. The visual information in V1 was 9.17×10^{-4} times that in the photoreceptor. Nevertheless, during the processing from V1 to V2, the dynamic degradation was scanty; the visual information in V2 was 3.74×10^{-1} times than that of V1.

3.1.4 Experiment of Mount Fuji

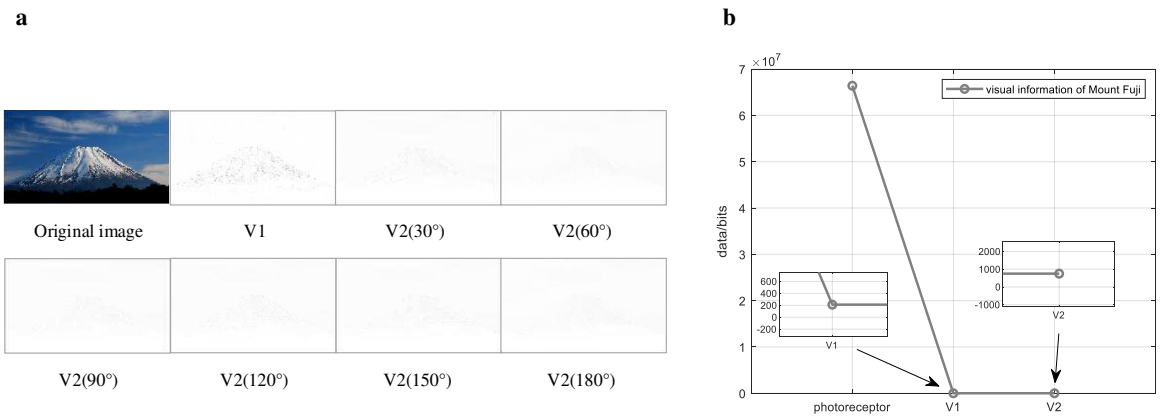

Fig. 12 *Mount Fuji image and PMVICV2 model responses.* **a.** Mount Fuji original image and the processed image in V1(60°) and V2 (30°, 60°, 90°, 120°, 150°, 180° with the same direction). **b.** The results of (a): dynamic degradation of Mount Fuji image processed by PMVICV2 model. The line that represents the visual information of V1 and V2 is very close to the x-axis. For more details, we have zoomed in to clarify.

As shown in **(a) of Figure 12**, the Mount Fuji image (original image), of which the resolution was 3840×2160, was used as the experimental object. The photoreceptor received the optical signal, of which the visual information was 6.64×10^7 bits. Subsequently, the visual information was processed by cones and rods, then passed to ganglion cells, and the edge features of the image can be detected. The dynamic degradation occurred after the visual information was transmitted to LGN. Simple cells with different preferred orientations are defined as $\theta_{60°,120°,180°}$ as well. The visual information in V1 was 22.4 bits， 26.9 bits and 5.83 $\times10^2$ bits, respectively, which was about 3.38×10^{-7} times, 4.05×10^{-7} times, and 8.79×10^{-6} times that of the photoreceptors. The visual information in V1 was obtained by the processing of RFs of ganglion cells and those of LGN. Finally, the visual information was transmitted from V1 to V2. The RFs in V2 had a strong response to the different corresponding angles and directions denoted as angle size, identifying the image feature information, as shown in Equation (14). The processed image is displayed in **(a) of Figure 12**. Lastly, the visual information in V2 is shown in Table 7:

Table 7 Visual information in V2 of the experiment of Mount Fuji (Unit: bits)

Orientation\Size	30°	60°	90°	120°	150°	180°
1	5.71×10^2	0.9	0.8	1.25	5.80×10^2	3.50×10^3
2	5.80×10^2	1.25	0.8	0.9	5.71×10^2	3.21×10^3
3	5.71×10^2	0.9	0.8	1.25	5.80×10^2	3.50×10^3
4	5.80×10^2	1.25	0.8	0.9	5.71×10^2	3.21×10^3
Average	5.76×10^2	1.075	0.8	1.075	5.76×10^2	3.35×10^3

The comparison of visual information between photoreceptors and V2 is shown in Table 8:

Table 8 Relationship between photoreceptors and V2 of the experiment of Mount Fuji

Orientation\Degraded rate\Size	30°	60°	90°	120°	150°	180°
1	8.61×10^{-6}	1.36×10^{-8}	1.21×10^{-8}	1.88×10^{-8}	8.74×10^{-6}	5.28×10^{-5}
2	8.74×10^{-6}	1.88×10^{-8}	1.21×10^{-8}	1.36×10^{-8}	8.61×10^{-6}	4.83×10^{-5}
3	8.61×10^{-6}	1.36×10^{-8}	1.21×10^{-8}	1.88×10^{-8}	8.74×10^{-6}	5.28×10^{-5}
4	8.74×10^{-6}	1.88×10^{-8}	1.21×10^{-8}	1.36×10^{-8}	8.61×10^{-6}	4.83×10^{-5}
Average	8.67×10^{-6}	1.62×10^{-8}	1.21×10^{-8}	1.62×10^{-8}	8.67×10^{-6}	5.06×10^{-5}

From the above analysis, drawing on the image of Mount Fuji, we have indicated that visual information changes from the retina to V1 and V2, as shown in **(b) of Figure 12**. It can be recognized that the average value of visual information of photoreceptors was 6.64×10^{7} bits; the average value of V1 was 2.11×10^{2} bits; the average value of V2 was 7.51×10^{2} bits. These values demonstrated that the visual information degraded significantly from photoreceptors to V1. The visual information in V1 was 3.18×10^{-6} times that in the photoreceptor. Nevertheless, during the processing from V1 to V2, the visual information in V2 remained constant, which was 3.57 times that of V1.

3.2 Results and Analyses

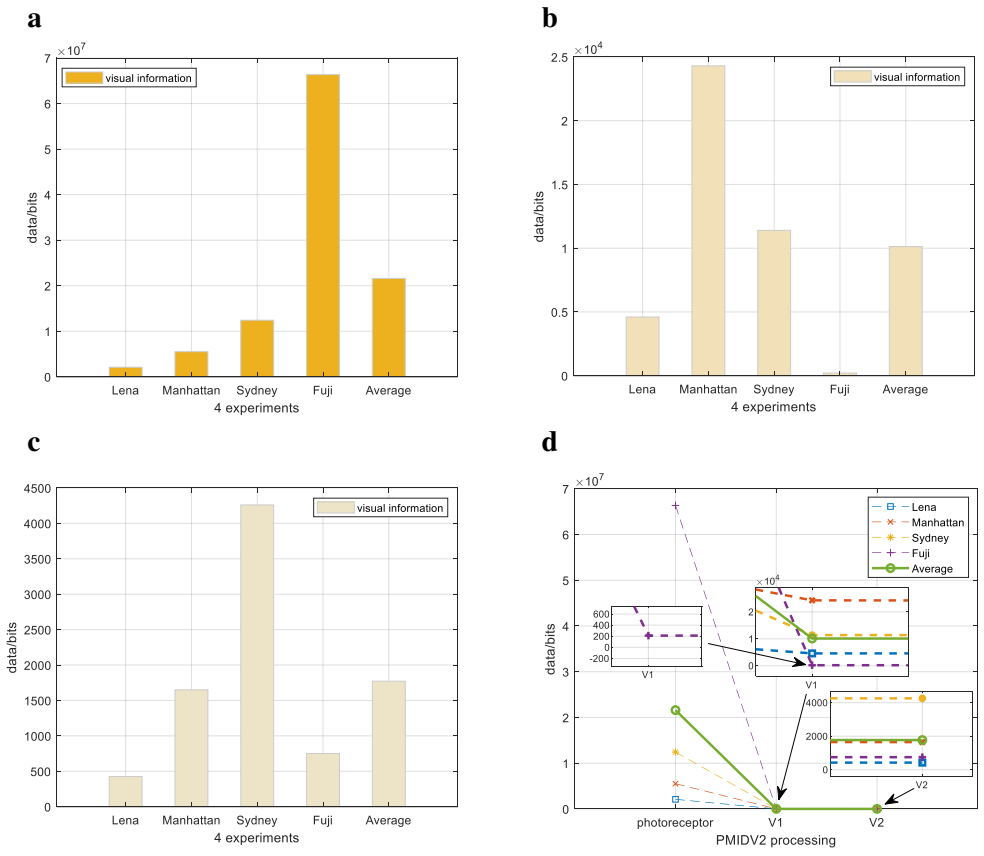

***Fig.13** Results.* **a.** Visual information of photoreceptors. **b.** Visual information of V1. **c.** Visual information of V2. **d.** According to (a), (b), and (c), visual information dynamic degradation of photoreceptor-V1-V2. The lines that represent the visual information of V1 and V2 are all overlapped and very close to the x-axis. For more details, we have zoomed in to clarify.

Table 9 Visual information changes in four experimental scenarios from PMVICV2 model

Scene\Data	Photoreceptor/bits	V1/bits	V2/bits	Photoreceptor to V2	V1 to V2
Lena	2.10×10^{6}	4.60×10^{3}	4.26×10^{2}	2.03×10^{-4}	9.25×10^{-2}
Manhattan	5.52×10^{6}	2.43×10^{4}	1.65×10^{3}	2.99×10^{-4}	6.77×10^{-2}
Sydney	1.24×10^{7}	1.14×10^{4}	4.26×10^{3}	3.43×10^{-4}	0.37
Fuji	6.64×10^{7}	2.11×10^{2}	7.51×10^{2}	1.13×10^{-5}	3.57
Average	2.16×10^{7}	1.01×10^{4}	1.77×10^{3}	8.19×10^{-5}	0.18

Based on the above experimental images, the visual information of photoreceptors from the PMVICV2 model was 2.10×10^{6} bits, 5.52×10^{6} bits, 1.24×10^{7} bits, and 6.64×10^{7} bits, respectively. The average value of those was 2.16×10^{7} bits, as shown in **(a) of Figure 13**.

After transmitting the visual information to the RFs of ganglion cells and LGN and V1 area, the data was calculated as 4.60×10^{3} bits, 2.43×10^{4} bits, 1.14×10^{4} bits, and 2.11×10^{2} bits, respectively. The average value was 1.01×10^{4} bits, which is shown in **(b) of Figure 13**.

Ultimately, the processed visual information was transmitted from V1 to V2, of which the value was 4.26×10^{2} bits, 1.65×10^{3} bits, 4.26×10^{3} bits, 7.51×10^{2} bits, respectively, and the average value was 1.77×10^{3} bits, shown in **(c) of Figure 13**.

Figure **(a)-(c) of Figure 13** showed that the visual information changes of the PMVICV2 model in these four scenarios could be obtained, as shown in **(d) of Figure 13** and Table 9. The visual information transmitted to V2 was 8.19×10^{-5} times that to photoreceptor and 0.18 times that to V1. Despite the different test images, there were no significant differences across the experimental results. It can be concluded that the significant dynamic degradation existed in the photoreceptor to V1 during the pathway of photoreceptor-ganglion cell-LGN-V1-V2. In the subsequent process of transmitting from V1 to V2, there had only a short dynamic degradation. Taken the analyses together, the significant dynamic degradation existed in the pathway of photoreceptor-ganglion cell-LGN-V1, which exhibited substantial differences between light and dark were retained by convolution calculation. Then, the edge signal of the image was obtained. In the process of visual information processing of the pathway of V1-V2, although the RFs in V2 had a strong response to the corner, they did not further extract the image feature, which accounted in part for the small dynamic degradation.

4 Conclusions

Taking into account energy metabolism, the brain capacity is actually limited in terms of fully transmitting visual information into the visual cortex, leading inevitably to visual information degradation. Then, how could the brain perceive the environment efficiently? Chumbley and Friston contend that surprise, captured by prediction error (defined as the difference between observed and expected quantities), drives learning (Chumbley et al., 2014; Friston, 2010). Our previous research shows one reason for degradation, which is related to prediction error, is that Retina-LGN-V1 has the convolution calculation, which acts to extract the pivotal visual information, ignore the unnecessary, and thus effectively saving brain power consumption. The findings serve as a further elaboration of the "prediction error" proposed by Friston. Building on this discovery, we are driven to further explore the visual information degradation or changes in

V1-V2. As a result, in undertaking this study, we seek to shed light on the mechanism by which the visual information is mapped from V1 to V2. Through establishing an original PMVICV2 model and conducting a quantitative analysis, it reaches four major conclusions stated as follows:

1) A quantitative description of visual information degradation in V1-V2.

According to the results of the PMVICV2 model, we can achieve Table 9, which shows the visual information in V2 is 8.19×10^{-5} times that of the photoreceptor and 0.18 times that of V1. It yields an exact quantitative interpretation of the visual information dynamic degradation in V2 by developing and experimenting with a new computational model. In doing so, it complements previous research wherein the neuroscientific experiment of the dynamic degradation focused chiefly on V1, which promotes a more accurate and specific understanding of the way visual information is encoded and managed in V2.

2) Strong response to the "corner" information, but a slight degradation

While moving from low-order to high-order visual signal processing, the visual information degrades significantly from the pathway of photoreceptor-ganglion cell-LGN-V1 (Raichle, 2010; Zhong & Wang, 2020). However, according to **(d) of Figure 13** and Table 9, the dynamic degradation has been scarcely observed during the mapping from V1 to V2. Whereas the RFs in V2 exhibit a strong response to the "corner" information (Hosoya & Hyvärinen, 2015), they do not further extract the image feature information. This demonstrates that a significant amount of dynamic degradation only has occurred on the pathway of photoreceptor-ganglion cell-LGN-V1, leaving limited visual information existing in V1 for the RFs in V2 to encode. This is a new discovery that has never been noticed before.

3) Convolution calculation in V1-V2

During the visual information processing (LeCun et al., 2015), the convolution calculation can be found on the pathway of photoreceptor-ganglion cell-LGN-V1 (Zhong & Wang, 2020). Moreover, the anatomical architecture between V1 and V2: one RF of V2 is weighted by two RFs of V1 (Hosoya & Hyvärinen, 2015), which suggests that the convolution calculation also exists in V1-V2.

4) STDP rule making more effective response to "corner" information

As we mentioned in Fig. 7, STDP rule intensifies the edge of the image and moderates the non-edge of the image. Therefore, the RFs of V2 can effectively respond to and encode "corner" information about the real world, dealing with the scarcity of visual information mapped from V1.

Despite the quantitative calculation and interpretation of the visual information changes in V1-V2, the study also has limitations. Structurally, we did not take all the details of Retina-LGN-V1-V2 into account due to the fact that the human visual system is complicated (see Figure 4), and that the visual information processing mechanisms have not been clearly uncovered (Raichle, 2010; Zhong & Wang, 2020). Therefore, we concentrated on the basic contour features such as edge and corner, which are considered highly relative to the degradation. Furthermore, we have not counted the top-down predictions since the novel visual information of the real world mapping from the retina to V2, which involves degradation, is a bottom-up transmission. According to Chumbley and Friston (Chumbley et al., 2014; Friston, 2010), bottom-up inputs make prediction errors, which originate from the novel visual information and are linked to degradation. The

mutual exchange of bottom-up prediction errors and top-down predictions from higher-order areas proceeds until prediction error is minimized. It means the degradation during the mapping from the retina to higher-order areas can be minimized likewise. This complex operative mechanism merits continued investigation in our future research.

Acknowledgments

This study was funded by the National Natural Science Foundation of China (Nos. 11472104, 11872180, 12072113).

Conflict of interest

All authors declare that they have no conflict of interest.

References

Amunts, K., Ebell, C., Muller, J., Telefont, M., Knoll, A., & Lippert, T. (2016). The human brain project: creating a European research infrastructure to decode the human brain. *Neuron, 92*(3), 574-581.

Anderson, C. H., Van Essen, D. C., & Olshausen, B. A. (2005). Directed visual attention and the dynamic control of information flow. In *Neurobiology of attention* (pp. 11-17): Elsevier.

Azzopardi, G., Rodríguez-Sánchez, A., Piater, J., & Petkov, N. (2014). A push-pull CORF model of a simple cell with antiphase inhibition improves SNR and contour detection. *Plos One, 9*(7).

Banich, M. T., & Compton, R. J. (2018). *Cognitive neuroscience*: Cambridge University Press.

Bargmann, C. I., & Newsome, W. T. (2014). The brain research through advancing innovative neurotechnologies (BRAIN) initiative and neurology. *JAMA neurology, 71*(6), 675-676.

Beyeler, M., Dutt, N. D., & Krichmar, J. L. (2013). Categorization and decision-making in a neurobiologically plausible spiking network using a STDP-like learning rule. *Neural Networks, 48*, 109-124.

Carver, S., Roth, E., Cowan, N. J., & Fortune, E. S. (2008). Synaptic plasticity can produce and enhance direction selectivity. *Plos Computational Biology, 4*(2), e32.

Chessa, M., Sabatini, S. P., & Solari, F. (2016). A systematic analysis of a V1–MT neural model for motion estimation. *Neurocomputing, 173*, 1811-1823.

Chumbley, J. R., Burke, C. J., Stephan, K. E., Friston, K. J., Tobler, P. N., & Fehr, E. (2014). Surprise beyond prediction error. *Human brain mapping, 35*(9), 4805-4814.

Cox, D., & Dean, T. (2014). Neural Networks and Neuroscience-Inspired Computer Vision. *Current biology : CB, 24*, R921-R929. doi:10.1016/j.cub.2014.08.026

Dubner, R., & Zeki, S. M. (1971). Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Research, 35*(2), 528-532.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience, 11*(2), 127-138.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2019). *Cognitive neuroscience: the biology of the mind*: W. W. Norton.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*: MIT press.

Hatori, Y., Mashita, T., & Sakai, K. (2016). Sparse coding generates curvature selectivity in V4 neurons. *JOSA A, 33*(4), 527-537.

Hosoya, H., & Hyvärinen, A. (2015). A hierarchical statistical model of natural images explains tuning properties in V2. *Journal of Neuroscience, 35*(29), 10412-10428.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology, 160*(1), 106-154.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194-203.

Joukes, J., Hartmann, T. S., & Krekelberg, B. (2014). Motion detection based on recurrent network dynamics. *Frontiers in Systems Neuroscience, 8*(7), 239.

Kendall, J. D., & Kumar, S. (2020). The building blocks of a brain-inspired computer. *Applied Physics Reviews, 7*(1), 011305. doi:10.1063/1.5129306

Khan, R. A., Meyer, A., Konik, H., & Bouakaz, S. (2012). *Human vision inspired framework for facial expressions recognition.* Paper presented at the 2012 19th IEEE international conference on image processing.

Kim, S.-Y., & Lim, W. (2019). Burst synchronization in a scale-free neuronal network with inhibitory spike-timing-dependent plasticity. *Cognitive neurodynamics, 13*(1), 53-73.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436.

Lee, H., Ekanadham, C., & Ng, A. Y. (2008). *Sparse deep belief net model for visual area V2.* Paper presented at the Advances in neural information processing systems.

Li, Z. (2003). V1 mechanisms and some figure-ground and border effects. *J Physiol Paris, 97*(4), 503-515.

Li, Z. (2014). *Understanding vision: theory, models, and data*: Oxford University Press, USA.

Liu, B.-h., Li, P., Sun, Y. J., Li, Y.-t., Zhang, L. I., & Tao, H. W. (2010). Intervening inhibition underlies simple-cell receptive field structure in visual cortex. *Nature Neuroscience, 13*(1), 89-96. doi:10.1038/nn.2443

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience, 7*(11), 3416-3468.

Marblestone, A. H., Wayne, G., & Kording, K. P. (2016). Toward an integration of deep learning and neuroscience. *Frontiers in computational neuroscience, 10*, 94.

Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information. In: San Francisco: WH Freeman.

Marr, D. (2010). *Vision: A computational investigation into the human representation and processing of visual information.*: Cambridge: MIT Press.

Meng, X., & Wang, Z. (2011). Enhancement Strategy for Region of Interest Based on Attentional Shroud. *Computer Engineering, 37*(4), 18-20.

Minami, I., & Naokazu, G. (2011). Mechanisms underlying the representation of angles embedded within contour stimuli in area V2 of macaque monkeys. *European Journal of Neuroscience, 33*(1), 130-142.

Obara, K., O'Hashi, K., & Tanifuji, M. (2017). Mechanisms for shaping receptive field in monkey area TE. *Journal of neurophysiology, 118*(4), 2448-2457.

Oberauer, K., & Lin, H. Y. (2017). An interference model of visual working memory. *Psychological Review, 124*(1), 21-59.

Okano, H., Miyawaki, A., & Kasai, K. (2015). Brain/MINDS: brain-mapping project in Japan. *Philosophical Transactions of the Royal Society B: Biological Sciences, 370*(1668), 20140310.

Olshausen, B., & Field, D. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature, 381*, 607-609. doi:10.1038/381607a0

Oprea, L., Pack, C. C., & Khadra, A. (2020). Machine classification of spatiotemporal patterns: automated parameter search in a rebounding spiking network. *Cognitive neurodynamics, 14*(3), 267-280. doi:10.1007/s11571-020-09568-8

Panetta, K., Gao, C., & Agaian, S. (2015). Human-visual-system-inspired underwater image quality measures. *IEEE Journal of Oceanic Engineering, 41*(3), 541-551.

Poo, M.-m., Du, J.-l., Ip, N. Y., Xiong, Z.-Q., Xu, B., & Tan, T. (2016). China brain project: basic neuroscience, brain diseases, and brain-inspired computing. *Neuron, 92*(3), 591-596.

Raichle, M. E. (2010). Two views of brain function. *Trends in cognitive sciences, 14*(4), 180-190.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience, 2*(11), 1019-1025.

Riesenhuber, M., & Poggio, T. (2002). Neural mechanisms of object recognition. *Current opinion in neurobiology, 12*(2), 162-168.

Riley, S. N., & Davies, J. (2020). A spiking neural network model of spatial and visual mental imagery. *Cognitive neurodynamics, 14*(2), 239-251. doi:10.1007/s11571-019-09566-5

Ringach, D. L. (2002). Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. *Journal of neurophysiology, 88*(1), 455.

Schölkopf, B., Platt, J., & Hofmann, T. (2006). Graph-Based Visual Saliency. *Proc of Neural Information Processing Systems, 19*, 545-552.

Semedo, J. D., Zandvakili, A., Machens, C. K., Byron, M. Y., & Kohn, A. (2019). Cortical areas interact through a communication subspace. *Neuron, 102*(1), 249-259. e244.

Shou, T. (2010a). *Brain mechanisms of visual information processing*. Hefei: University of Science and Technology of China Press.

Shou, T. (2010b). The functional roles of feedback projections in the visual system. *Neuroscience Bulletin, 26*(5), 401-410.

Ungerleider, L. G., & Haxby, J. V. (1994). 'What' and 'where' in the human brain. *Current opinion in neurobiology, 4*(2), 157-165.

Xiao, J., & Huang, X. (2015). Distributed and dynamic neural encoding of multiple motion directions of transparently moving stimuli in cortical area MT. *Journal of Neuroscience, 35*(49), 16180-16198.

Xu, Y., Lv, C., Li, S., Xin, P., Ma, S., Zou, H., & Zhang, W. (2017). Review of development of visual neural computing. *Computer Engineering & Applications*.

Yin, Y., Li, Z., & Wu, L. (2018). Bottom-up saliency and top-down learning in the primary visual cortex of monkeys. *Proceedings of the National Academy of Sciences, 115*(41), 10499-10504.

Yu, Z., Liu, J. K., Jia, S., Zhang, Y., Zheng, Y., Tian, Y., & Huang, T. (2020). Toward the next generation of retinal neuroprosthesis: visual computation with spikes. *Engineering, 6*(4), 449-461. doi:https://doi.org/10.1016/j.eng.2020.02.004

Zhao, S., Zou, Q., Jin, Z., Yao, G., & Li, Y. (2010). Neural computation of visual imaging based on Kronecker product in the primary visual cortex. *BMC Neuroscience,11,1(2010-03-26), 11*(1), 1-13.

Zhaoping, L. (2019). A new framework for understanding vision from the perspective of the primary visual cortex. *Current opinion in neurobiology, 58*, 1-10.

Zhong, H., & Wang, R. (2020). Neural mechanism of visual information degradation from retina to V1 area. *Cognitive neurodynamics*. doi:10.1007/s11571-020-09599-1

Ziemba, C. M., Freeman, J., Simoncelli, E. P., & Movshon, J. A. (2018). Contextual modulation of sensitivity to naturalistic image structure in macaque V2. *Journal of neurophysiology, 120*(2), 409-420.

Figures

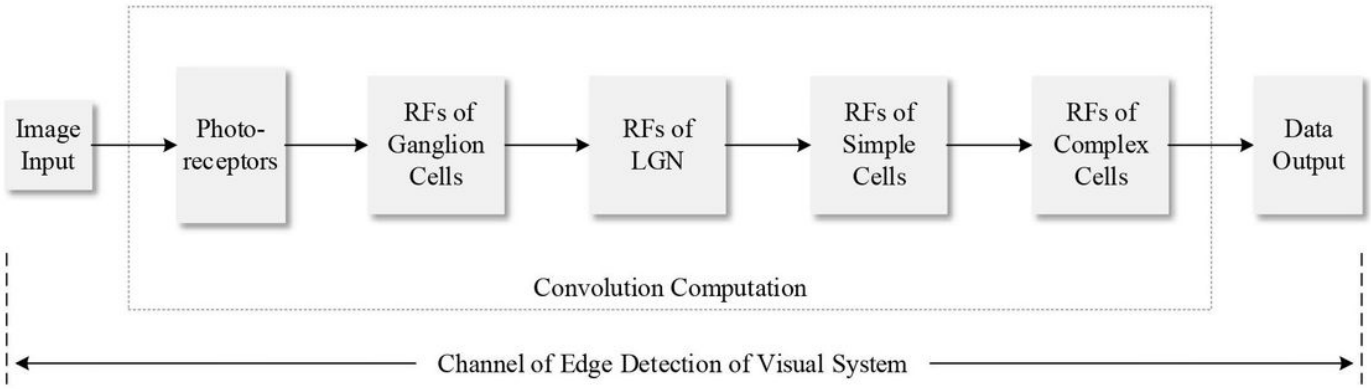

Figure 1

The visual information processing from retina to V1.

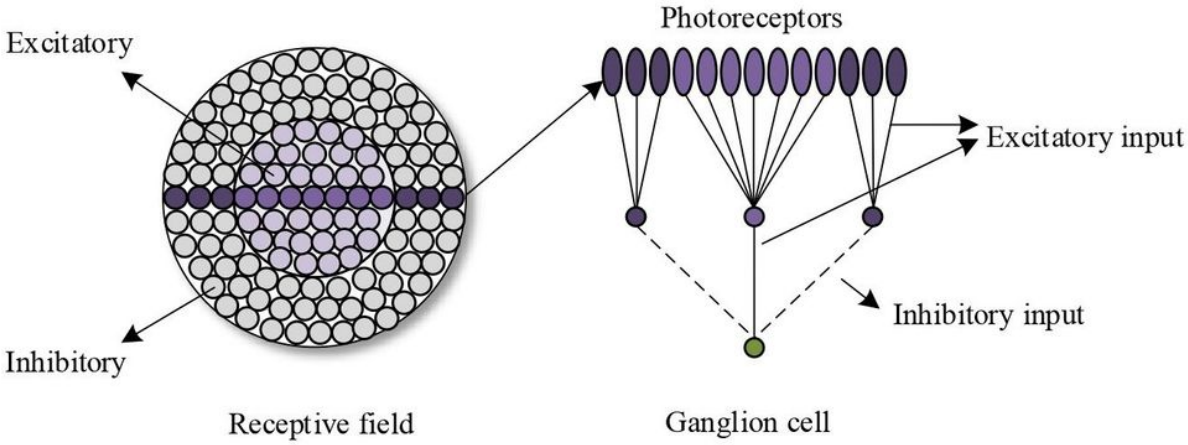

Figure 2

On-center model of ganglion cell.

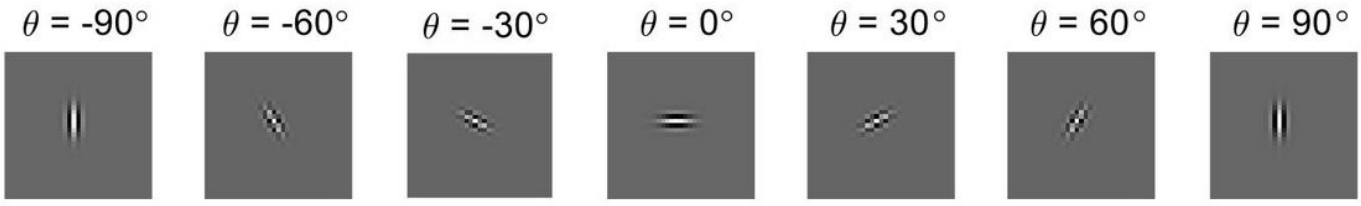

Figure 3

Simple cells for different preferred orientations.

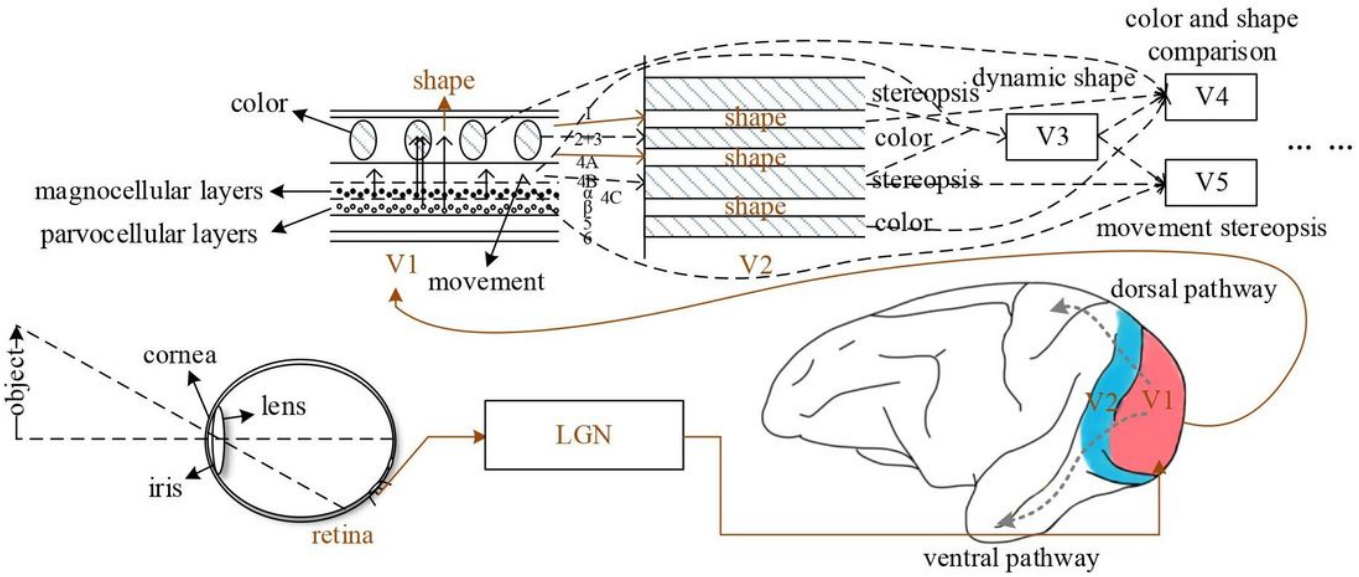

Figure 4

The diagram of visual information processing and transmission.

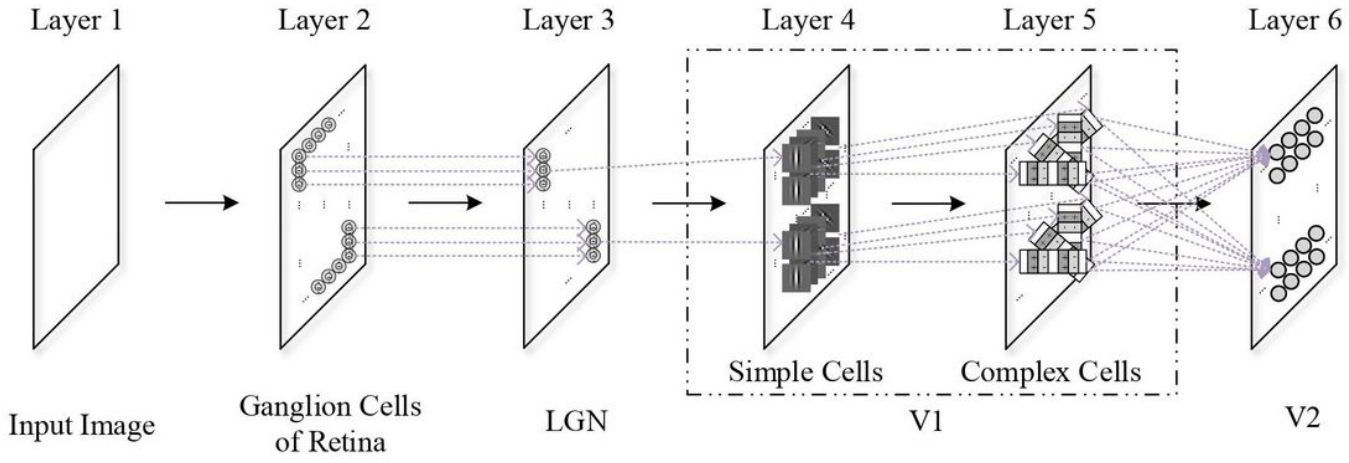

Figure 5

The structural diagram of visual information from retina to V2.

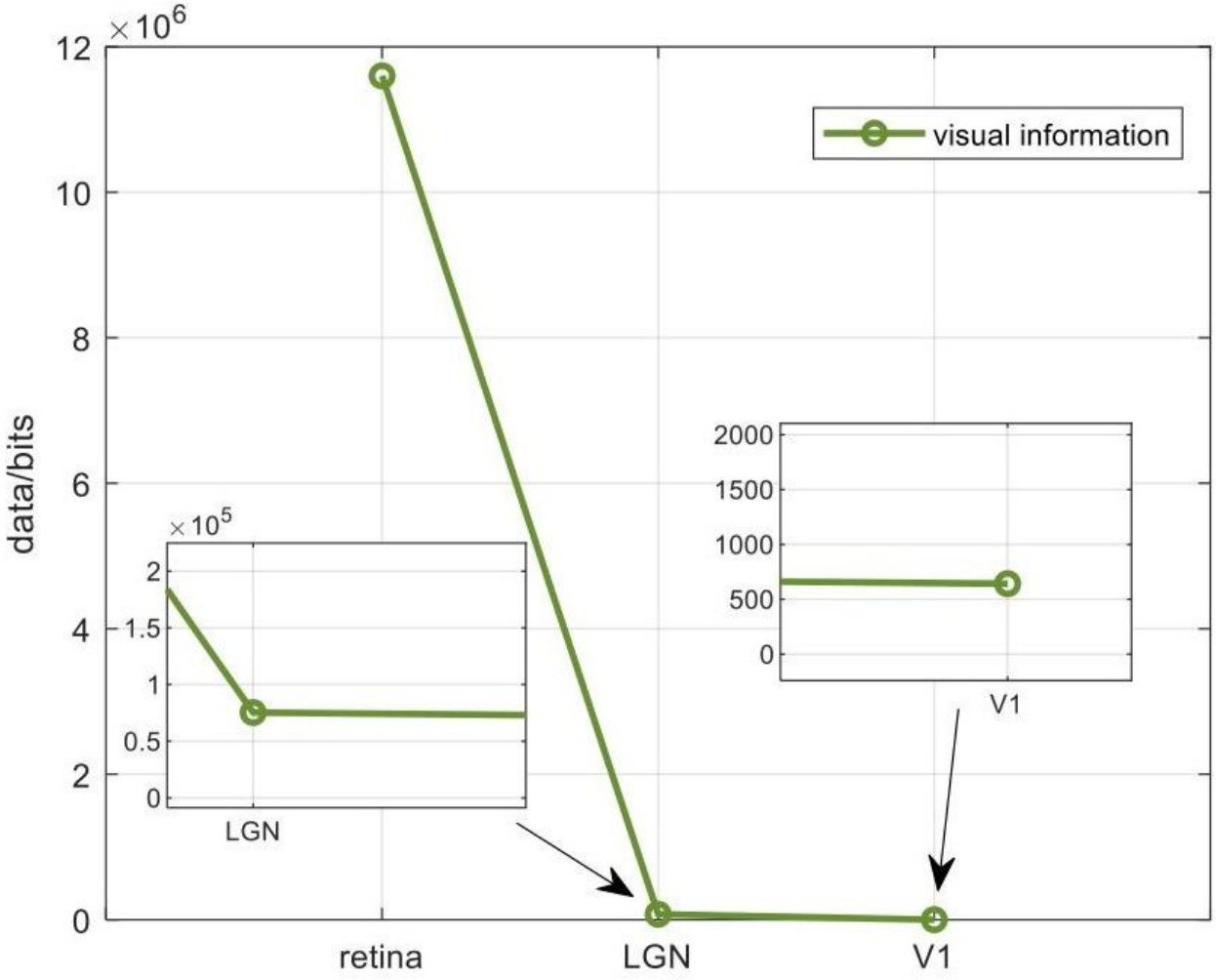

Figure 6

The visual information dynamic degradation of photoreceptor-ganglion cell-LGN-V1 based on EDMRV1 model (Zhong & Wang, 2020). The line that represents the visual information of LGN and V1 is very close to the x-axis. For more details, we have zoomed in to clarify.

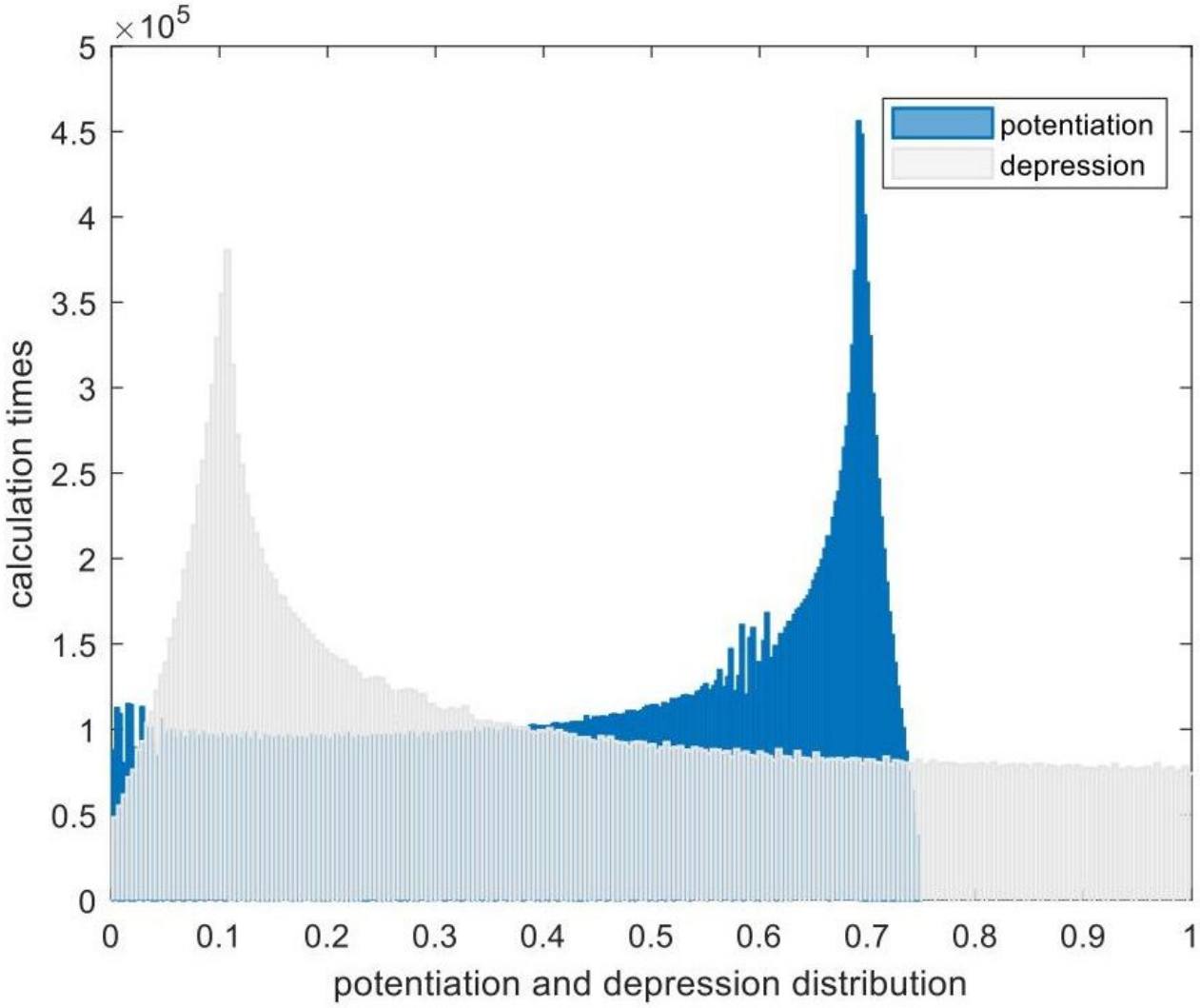

Figure 7

STDP rule. The potentiation and depression distribution of weights. The x-axis indicates the results of potentiation(i, j) and depression(i, j); the y-axis, the calculation times, which means the higher value of the y-axis indicates that the higher frequency of appearance of the corresponding x value.

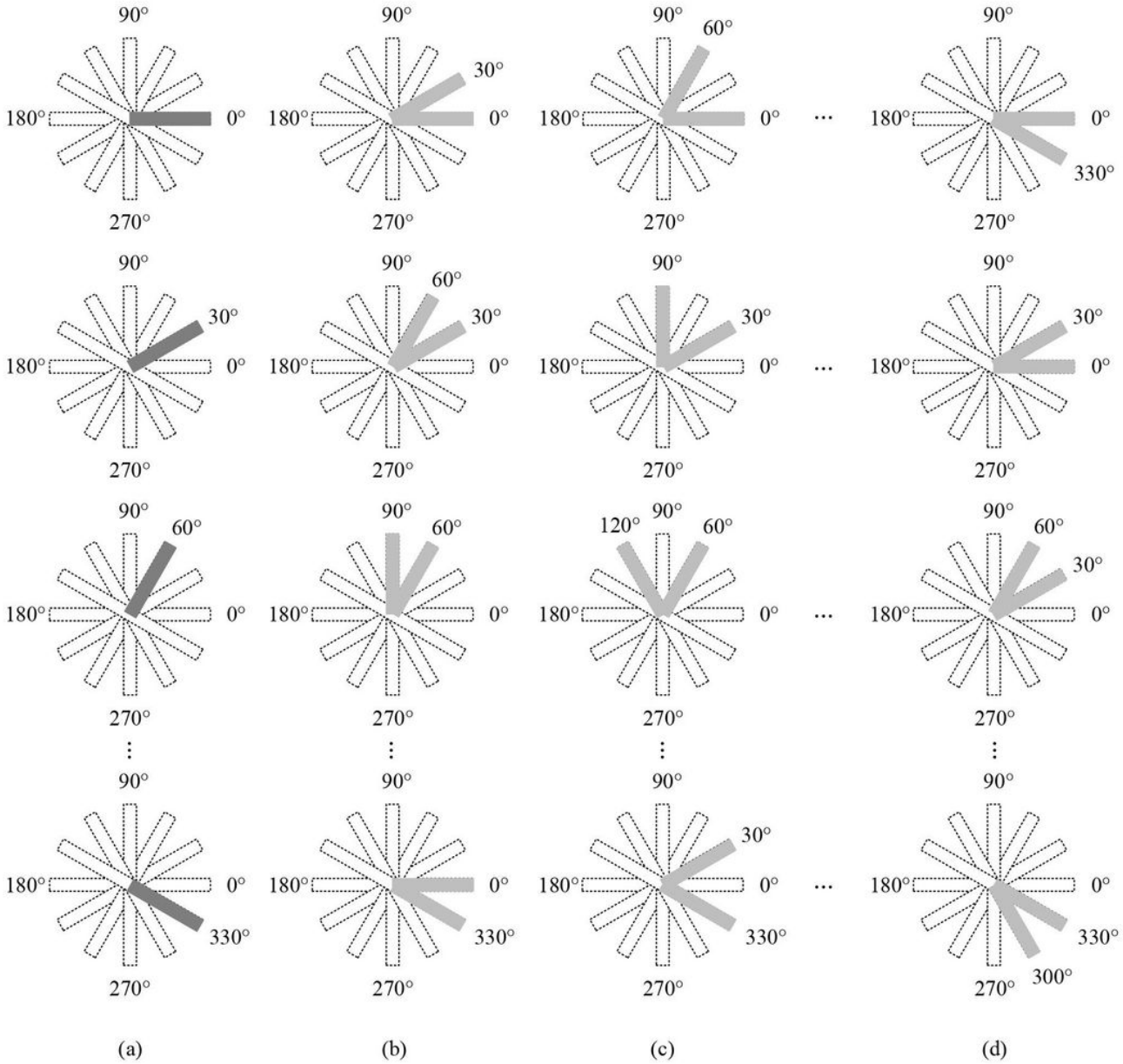

Figure 8

The RFs (two grey sides) in V2 with varying angles and directions. a. The angles of RFs equal 0° in different directions. The dark grey indicates two sides of RFs are overlapped. b. The angles of RFs equal 30° in different directions. c. The angles of RFs equal 60° in different directions. d. The angles of RFs equal 330° in different directions.

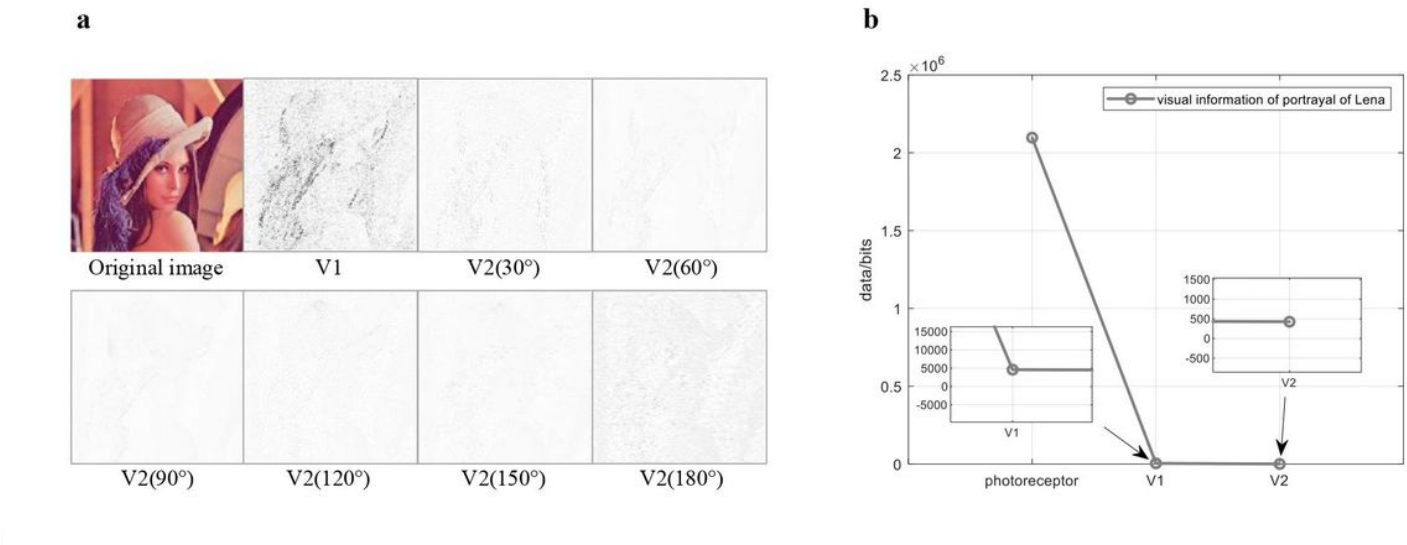

Figure 9

Lina image and PMVICV2 model responses. a. Lina original image and the processed image in V1(60°) and V2 (30°, 60°, 90°, 120°, 150°, 180° with the same direction). b. The results of (a): dynamic degradation of Lina image processed by PMVICV2 model. The line that represents the visual information of V1 and V2 is very close to the x-axis. For more details, we have zoomed in to clarify.

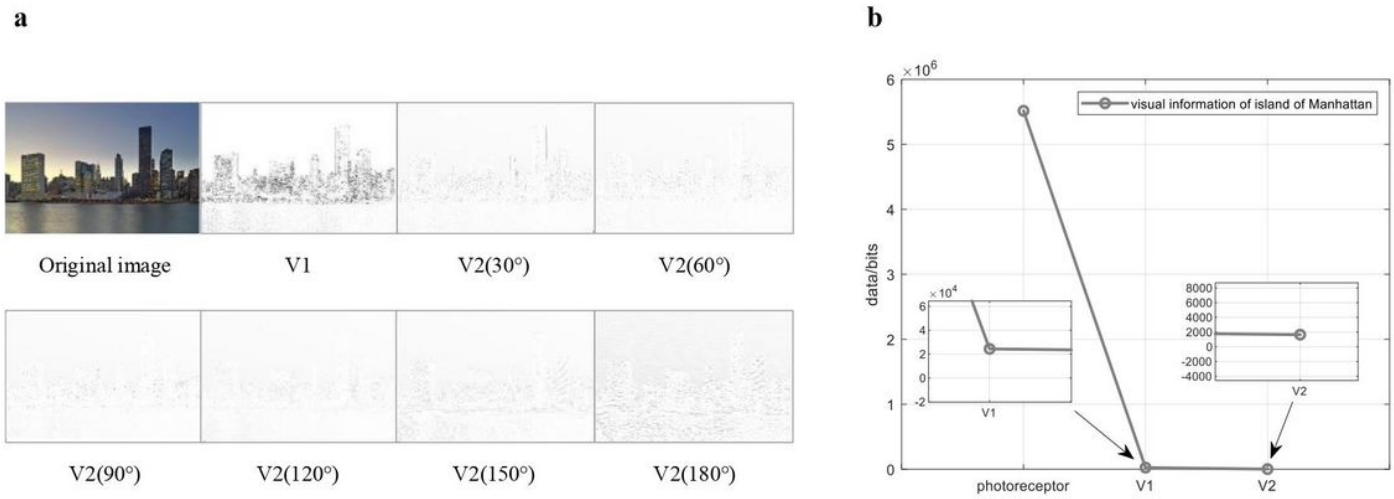

Figure 10

Manhattan image and PMVICV2 model responses. a. Manhattan original image and the processed image in V1(60°) and V2 (30°, 60°, 90°, 120°, 150°, 180° with the same direction). b. The results of (a): dynamic degradation of Manhattan image processed by PMVICV2 model. The line that represents the visual information of V1 and V2 is very close to the x-axis. For more details, we have zoomed in to clarify.

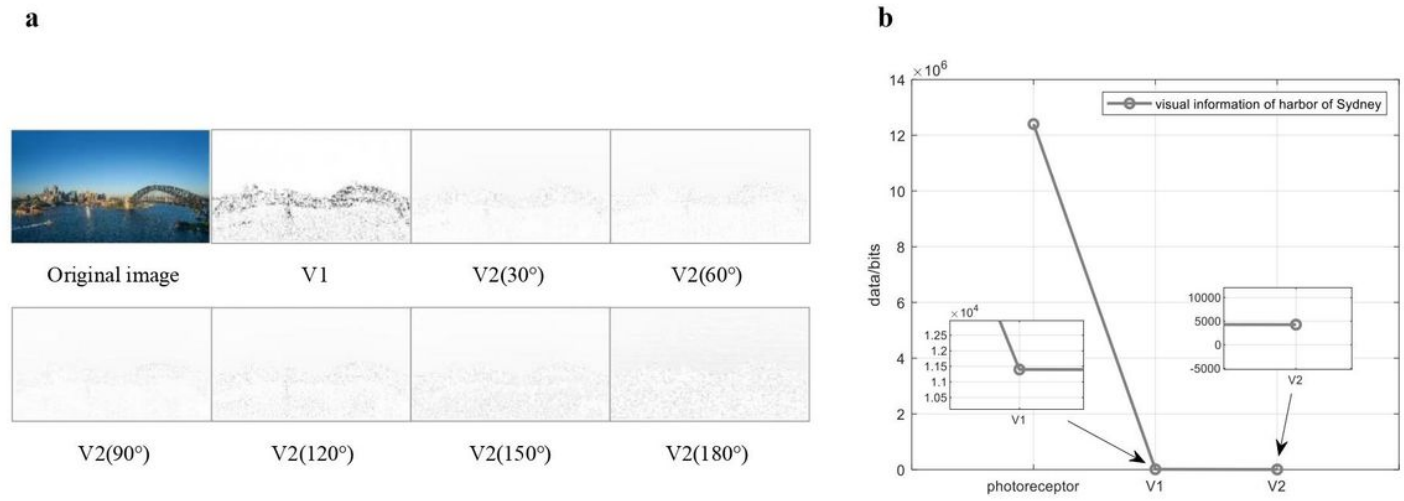

Figure 11

Sydney image and PMVICV2 model responses. a. Sydney original image and the processed image in V1(60°) and V2 (30°, 60°, 90°, 120°, 150°, 180° with the same direction). b. The results of (a): dynamic degradation of Sydney image processed by PMVICV2 model. The line that represents the visual information of V1 and V2 is very close to the x-axis. For more details, we have zoomed in to clarify.

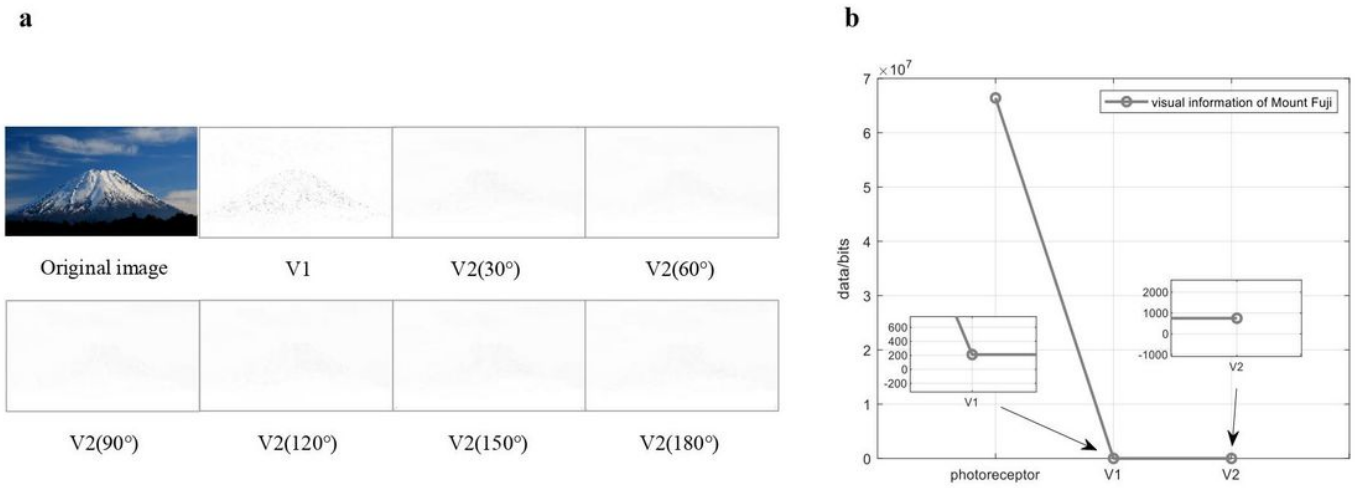

Figure 12

Mount Fuji image and PMVICV2 model responses. a. Mount Fuji original image and the processed image in V1(60°) and V2 (30°, 60°, 90°, 120°, 150°, 180° with the same direction). b. The results of (a): dynamic degradation of Mount Fuji image processed by PMVICV2 model. The line that represents the visual information of V1 and V2 is very close to the x-axis. For more details, we have zoomed in to clarify.

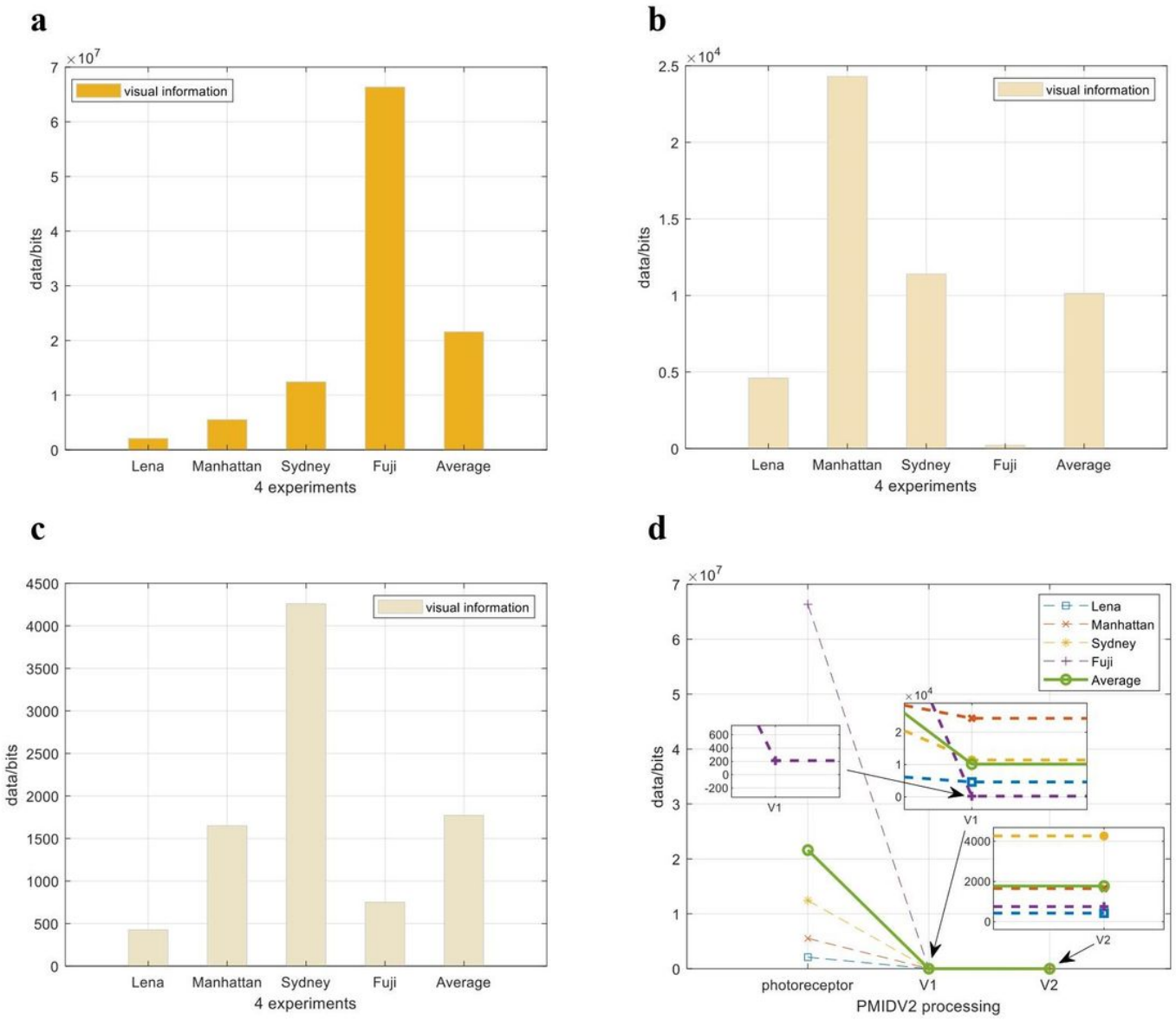

Figure 13

Results. a. Visual information of photoreceptors. b. Visual information of V1. c. Visual information of V2. d. According to (a), (b), and (c), visual information dynamic degradation of photoreceptor-V1-V2. The lines that represent the visual information of V1 and V2 are all overlapped and very close to the x-axis. For more details, we have zoomed in to clarify.