Seeing the world in three dimensions, fast-but-rough and slow-but-accurate

When we perform a task, there is generally the tradeoff between the time taken to achieve the goal and the quality of the attained result. For example, when we solve a mathematical problem, we can quickly perform the calculation at the cost of potential errors, or we can make sure that the result is accurate by spending more time. We often face a similar tradeoff in our daily life. Naturally, an efficient way to perform a task is to switch between the time-efficient and quality-guaranteed strategies depending on the situation at hand. For some tasks, even more appealing solution would be to apply the two strategies in parallel and combine their results flexibly according to the current need. We propose that our visual system may implement this kind of flexible mechanism to perform an important visual task: to see the world in three dimensions, or stereopsis.
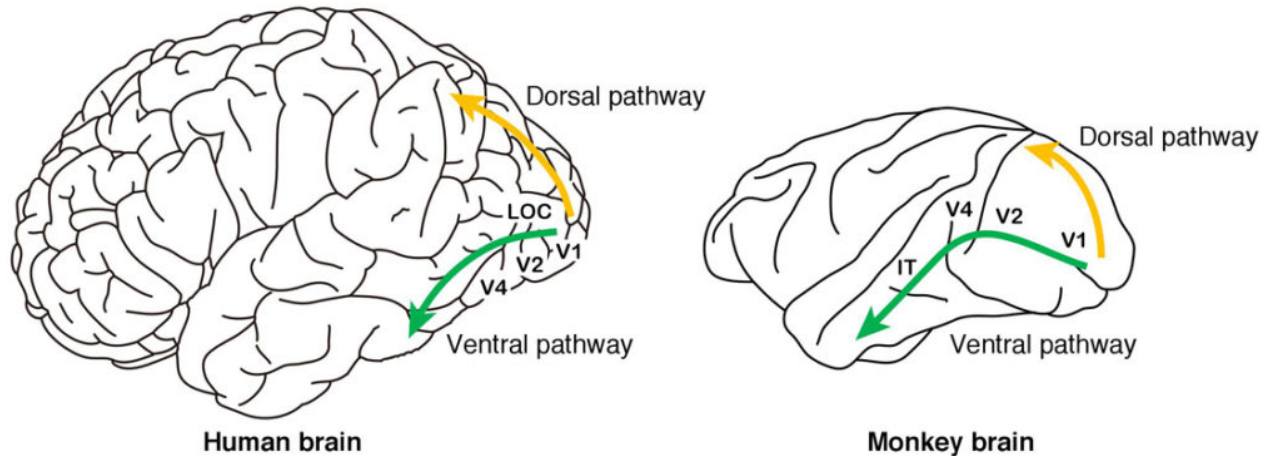

Fig. 1. Binocular disparity is detected by the primary visual cortex (V1) and further processed by the ventral pathway projecting to the temporal cortex as well as the dorsal pathway projecting to the parietal cortex. One of the visual areas along the ventral pathway, V4, plays a critical role in slow but accurate sense of depth.

Binocular disparity, a slight difference between our left-eye and right-eye images, is a powerful cue for the three-dimensional visual perception. The initial encoding of binocular disparity takes place in the primary visual cortex (V1) and uses a relatively simple computation similar to cross-correlation (correlation-based computation); the inputs from the left eye and right eyes are multiplied with varying offsets to find the binocular disparity. The result of correlation computation is passed onto both of the two major visual cortical pathways, i.e., ventral and dorsal pathways (Fig. 1). The dorsal pathway largely inherits the result of correlation computation, whereas the ventral pathway further

refines the disparity signal using sophisticated mechanisms. The signal refinement removes the false disparity signals from the spurious matches between the left-eye and right-eye images and extracts the valid disparity signals for the true matches (match-based computation).

We created a new set of visual stimuli to estimate the relative weights of the correlation-based computation and match-based computation in the perceptual judgments of stereoscopic depth by human observers. We found that a simple, correlation computation contributed more when the task was to judge depth coarsely. A more sophisticated, match-based computation was recruited when the task was to judge fine depth. Also, the relative weights varied depending on a temporal nature of visual stimuli. The contribution of the correlation computation increased when subjects looked at a more dynamic, rapidly changing stimuli, whereas the contribution of the match-based computation increased for more slowly changing stimuli. A computational model in which both computations were combined with adaptive weights explained the observed results well. Thus, these psychophysical experiments revealed remarkable flexibility of our depth perception. Simple and complex mechanisms are adaptively recruited depending on the spatial resolution and temporal property of the input images.

We also analyzed the activity of visual cortical neurons of the monkeys looking at the same set of visual stimuli. These experiments revealed how the visual cortex performed the match-based computation. We recorded from a mid-tier cortical area along the ventral pathway called V4. When the activity of single V4 neurons was compared to the activity of V1 neurons, the difference was not striking. However, when we pool the recorded activity of many V4 neurons just in the way our visual system does, we found a clear advancement in the quality of the disparity signals in V4. These results suggest that pooling the activity of V4 neurons plays an important role in the neural mechanisms of stereoscopic depth perception.

Takahiro Doi[1]**, Ichiro Fujita**[2]
[1]University of Pennsylvania, Philadelphia, USA
[2]Osaka University, Suita, Japan

Publications

Weighted parallel contributions of binocular correlation and match signals to conscious perception of depth.
Fujita I, Doi T
Philos Trans R Soc Lond B Biol Sci. 2016 Jun 19

Pooled, but not single-neuron, responses in macaque V4 represent a solution to the stereo correspondence problem.
Abdolrahmani M, Doi T, Shiozaki HM, Fujita I