A Novel Illusion Involving Complex Motion and MT-MST Cortical Interaction

Masataka Watanabe [1], Ryou Ishizaki [1] and Kazuyuki Aihara [2]

METR 2002-11 January 2003

1 Department of Quantum Engineering and Systems Science, Graduate School of Engineering, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8656, Japan;

2 Department of Mathematical Engineering and Information Physics, Graduate School of Engineering, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8656, Japan;

Abstract

Growing evidence from psychophysics and single-unit recordings suggests specialized mechanisms in the primate visual system for the detection of complex motion patterns such as expansion and rotation. Here we used a complex motion simulating technique to determine the properties of these complex motion templates, namely, how we perceive visual input subtly different from the original pattern. For expansion, contraction and rotation, we found that the small difference is enhanced so as to perceive away from the perfect complex motion. We suggest the possibility of an inhibitory top down projection from the MST to the MT area which may explain the novel illusion.

1 Introduction

It is generally believed that neural processing of sensory information follows a hierarchical structure and proceeds along multiple parallel pathways. For instance, visual sensory input first enters the retina and is sent to the primary visual cortex through the LGN. Visual information is next passed on to area V2, where processing is divided into two pathways, namely, the dorsal pathway for motion and position processing which includes MT and MST areas, and the ventral pathway for color and shape processing which includes V4, TEO and TE areas. Physiological evidence is mounting that neurons at each ensuing area respond to more and more complex visual stimuli.

Fundamental work by Hubel and Wiesel has demonstrated that neurons in the primary visual cortex are characterized by well-defined tuning properties and the neurons are cyclopaedically available. That is to say, the dimensionality of the feature space is fairly low and there are enough neurons to cover the entire feature space with high resolution. The complexity of the tuning properties increases along the visual pathway which can be accounted for by the assumption that complex feature detectors are formed by combining the output signals from simple feature detectors on earlier levels. This combinatorial process leads to rapid increase of feature space dimensionality. This problem is known as the “curse of dimensionality”: the neurons that is required to encode a multi-dimensional feature space, assuming a fixed resolution of representation, increases exponentially with the dimensionality of feature space (Girosi et al.). One way to cope with this problem is to control the resolution of representation (Vapnik). We can actually see such control of resolution in the visual cortex, that is to say, the precision with which the retinal position is encoded in the primary visual cortex (areas V1 and V2) does not seem to be retained in the motion sensitive area MT (Gatass et al.). Another strategy which the brain might take in solving the dimensionality problem is to put a constraint on the coverage of the feature space reflecting the properties of objects in the natural environment. Objects in the outer world are not completely arbitrary and random but tend to be continuous, piece-wise smooth, symmetric and usually familiar to the observer. In area TE, the final area of the ventral pathway, neurons are selective for complex shapes and estimates say that there are 1300-2000 columns in a single hemisphere. This estimated number of columns should be fairly small compared to the high dimensionality of the feature space and only a portion must be covered. These covered “islands” may be defined as “templates”.

We are interested in the properties of the templates in the high dimensional feature space. How do we perceive when given a visual input slightly different from a particular template? Is it perceived closer to or away from the template, in other words, is the slight difference attenuated or emphasized? To answer this fundamental question, we take up the complex motion template which is related to the perception of self-motion found in the final area of the dorsal visual pathway, the MST area.

It is known that important source of self-motion perception comes from visual information (Gibson). As we move through the environment, the pattern of visual motion on the retina, known as “optical flow” provides rich information about our self-motion. Moving forward in space is combined with “expansion”, moving backward with “contraction” and turning of the head with “rotation” optical flows.

The reason we take up the complex motion template is that it is based on optical flow accompanied by self-motion and the properties of these templates are well defined. They are perfect rotation, perfect expansion, perfect contraction and linear combinations of them. We performed experiments using a set of random dot motion patterns which simulated rotation, expansion and contraction and focused on how they were perceived compared to the perfect complex motion. Our results were analyzed and were taken into consideration with the neural properties and MT-MST cortical interactions.

2 General Methods

2.1 Stimuli

Random dot motion stimuli were generated using a VSG 2/3 graphic board (Cambridge Research Systems) controlled by a PC computer and presented on a 17 inch color Trinitron monitor. In all cases, the random dots were plotted into an 800x800 pixel square region of size 17.1 deg x17.1 deg viewing angle.

Each dot consisted of a square of four white pixels (50cd/m2) which extended over a visual angle of 0.11 deg against a black background. The fixation point was a filled circle of diameter 0.22 deg.

The square area of stimuli was divided into four regions and each quadrant had independent rules for random dot motion. The rules were decided in a way to mimic complex motion for the test stimulus, namely, rotation for experiment 1 and expansion and contraction for experiment 2. The control stimulus was composed of the same set of local random dot motions but with a different arrangement. The speed of random dot motion was set at 12.1 deg/sec at the center of the quadrant. For all experiments, 400 dots in the stimulus square were presented at any moment.

2.2 Procedure

For all experiments, subjects performed a two-alternative, forced-choice discrimination task on a set of random dot motion. The experiments were conducted in a darkened room. After 500 ms delay from fixation onset, the random dot motion stimulus was presented for 500 ms while the subjects fixated on the fixation spot. Subsequent to the presentation of stimuli, subjects were required to discriminate between two categories, "incurved" or "excurved" for experiment 1, "expanding" or "contracting" along the direction of local random dot motion for experiment 2. The results were analyzed by calculating percent perceived values of the first category and plotting psychometric functions for each random dot motion type.

2.3 Observers

One of the author (RI) and two naive subjects (YF and TM) participated in the experiments. All had normal or corrected visual acuity.

3 Experiment 1

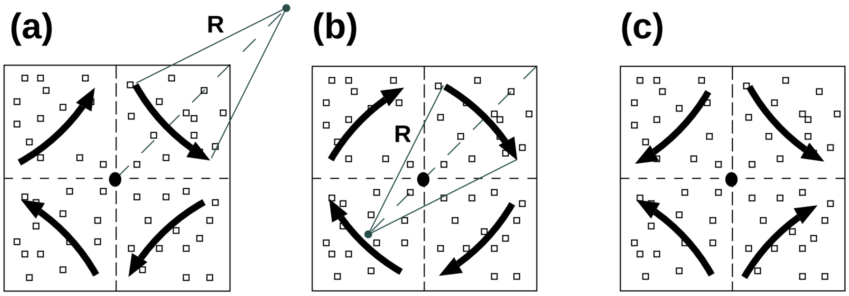

Figure 1: Schematic illustration of visual stimuli used in experiment 1. (a) incurved test stimulus. (b) excurved test stimulus. (c) incurved control stimulus.

The first experiment concentrated on the attracting or repelling properties of the complex motion template, rotation. The visual stimuli were divided into four quadrants with random dot rotation of uniform radius and independent center of rotation (fig. 1). The center of rotation of each quadrant was aligned on the diagonal line of the quadrant which included the center of the visual stimulus. When the center of rotation was put on the same side as the center of the visual stimulus, the whole stimuli became "incurved" (fig. 1 (a)). To the contrary, when the center of rotation was put on the opposite side, the whole stimuli became "excurved" (fig. 1 (b)). As for the direction of rotation, uniform direction, either clockwise or counterclockwise was used for all four quadrants in the test stimuli, and mixed direction was used for the control stimuli (fig. 1 (c)).

Subjects were required to perform a 2AFC discrimination task, whether the stimulus appeared "incurved" or "excurved". The order of the control and test stimulus was randomized and each particular movie was shown 10 times. Figure 2 shows the frequency at which the subjects reported the stimulus to be "incurved" plotted as a function of the actual curvature. Here, the x-axis represents the actual radius of random dot circular motion at the center of the quadrants given as 1/logR for incurved stimuli and -1/logR for excurved stimuli, whereas R is the radius. The value zero represents perfect straight local motion where R is infinitely large.

All three subjects indicated a clear trend that the perceived curvature for test stimulus lean to the "incurved" side, while no such bias existed for the control stimuli (fig.2). It can be concluded that the complex motion template repels the visual stimuli which mimic rotation and the difference between perfect rotations is enhanced.

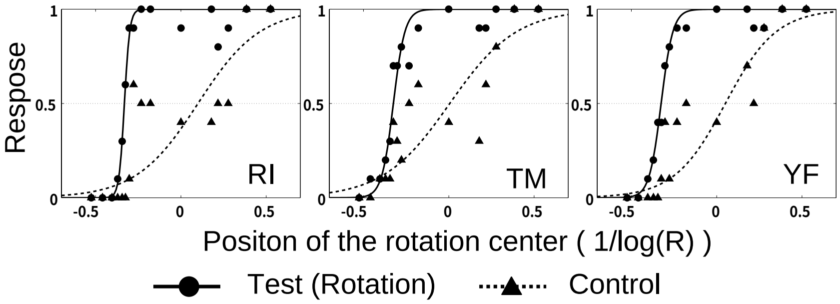

Figure 2: Individual subjects' data from experiment 1. The x-axis represents the actual curvature of local random dot motion in the quadrants. The y-axis represents the fraction of trials in which the visual stimuli is judged incurved. Each plot shows the psychophysical performance curve for a different observer. In each case, the point of perceptual equivalency is shifted to the left for test stimuli, indicating that each subject tended to judge the rotation simulating visual stimuli incurved.

4 Experiment 2

In the next experiment we tested the properties of expansion and contraction complex motion template. The local random dot motion in the four quadrants is expansion or contraction with the center of focus aligned on the diagonal line of the quadrants including the center of the visual stimulus (fig. 3). For visual stimuli mimicking global expansion, dot motion is said to be "expanding" when the center of focus was set on the same side as the center of the visual stimuli along the diagonal line of the quadrant and the random dots are expanding from it (fig. 3 (a)). On the other hand, dot motion is said to be "contracting" when the center of focus is on the opposite side to the center of the visual stimulus and the random dots are contracting to it (fig. 3 (b)). Other way around for the visual stimuli

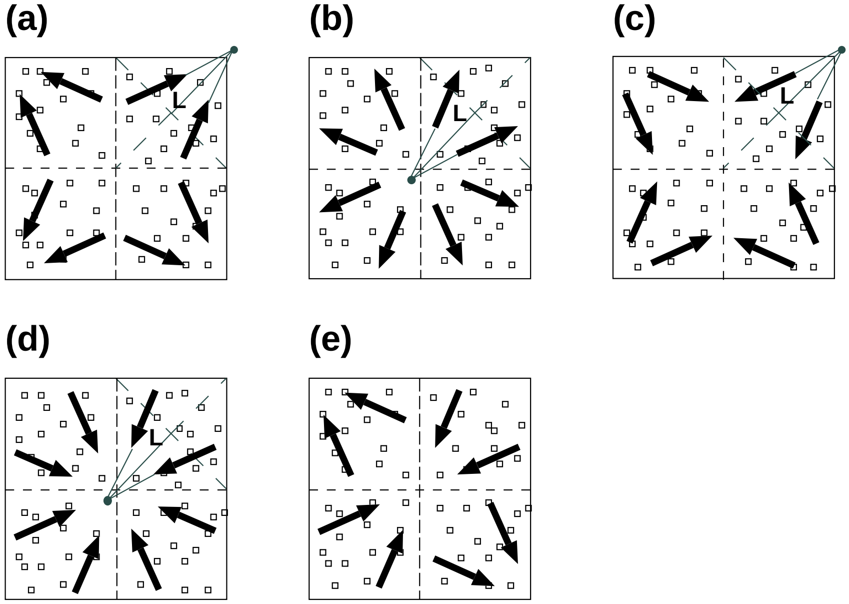

Figure 3: Schematic illustration of visual stimuli used in experiment 2. (a) locally contracting stimulus for global expansion. (b) locally expanding stimulus for global expansion. (c) locally expanding stimulus for global contraction. (d) locally contracting stimulus for global contraction. (e) locally contracting control stimulus.

which mimic global contraction (fig. 3 (c), (d)). For control stimuli, dot motions were mixed and placed diagonally (fig. 3 (e)). As in the first experiment, subjects judged whether the random dot motion appeared expanding or contracting within the four quadrants (2AFC task). The x axis gives 1/L, where L is the distance between the focus point of random dot motion and the center of the quadrant (fig. 4). For stimuli simulating global contraction, 1/L for contracting random dot motion and -1/L for expanding motion. To the contrary, in expansion simulating stimulus, 1/L for expanding motion and -1/L for contracting motion. The y axis gives the percentage over ten trials of subjects responding, “contracting” for expansion simulating stimulus and ”expanding” for contraction simulating and control stimulus. For test stimuli, all three subjects perceived so that local motion of random dots are pushed over to the opposite side of the global stimulus where no such bias existed for control stimuli.

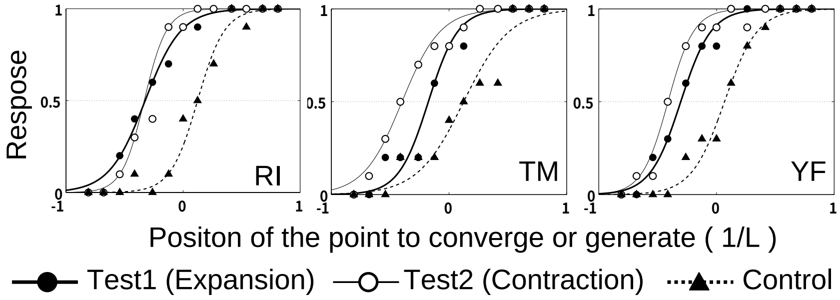

Figure 4: Individual subjects’ data from experiment 2. The x - axis represents the actual intensity of local expansion or local contraction within the quadrants. The y-axis represents the fraction of trials in which the visual stimuli is judged contracting for stimuli simulating global expansion and fraction of trials judged expanding for stimuli simulating global contraction. Each plot shows the psychophysical performance curve for a different observer. In each case, the point of perceptual equivalency is shifted to the left for test stimuli, indicating that each subject tended to judge the visual stimuli away from the global complex motion template.

5 Discussion and Conclusion

In both experiments, all subjects perceived the rotation, contraction and expansion simulating stimulus to deform away from perfect complex motion. In this section, we will take into consideration the neural mechanism of the novel illusion.

Cortical area MSTd is known to be the prime candidate for the computation of optical flow. MSTd neurons respond selectively to global visual patterns of rotation, expansion, contractions and combinations of them with large receptive fields, while neurons in other cortical areas related to visual motion such as V1 or MT have much smaller receptive fields and respond relatively to simple motions in a single direction (Saito et al.).

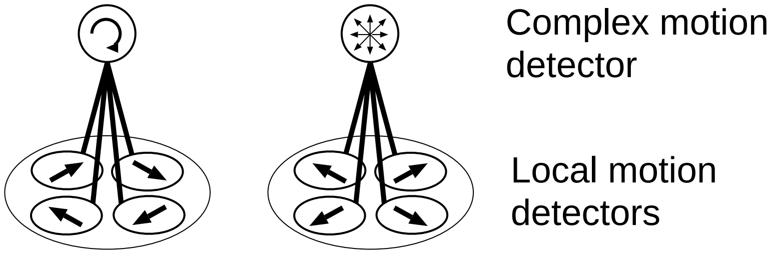

Figure 5: Model for complex motion detection. Projections of local motion detectors aligned spatially converge to the complex motion detector.

The basic explanation on how MSTd neurons analyze optical flow takes in account projections from velocity selective neurons in the MT region. The template model (fig. 5) assumes that a MSTd neuron selective to a particular complex motion trajectory integrates input from an array of MT cells with corresponding motion selectivity and receptive field (Saito et al.). Perrone et al. extended the model and verified it using physiological data. Moreover, Morrone et al. evaluated the S/N threshold in perceiving complex motions by adding noise to the basic optical flow stimuli with results supporting the existence of template mechanism in the brain.

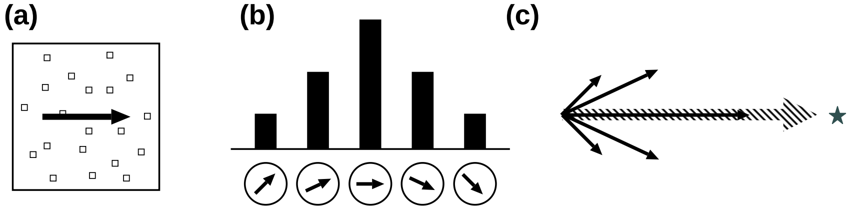

Figure 6: Population vector coding for local motion detection. (a) direction of random dot motion. (b) activity of neurons with direction tuning. (c) perceived direction given as a population vector of the activity of direction tuned neurons.

To explain the neural mechanism of our novel illusion, we bring in the effect of top down signals to the conventional template model. The role of top down signals in the nervous system remains controversial and there are two major hypotheses today. The first assumes a positive feedback for increased basin of attraction (ART: Grossberg) or increased speed of neural processing. On the other hand, the second hypothesis assumes a negative feedback, where the output of the lower module reflects the error between actual visual input and prediction of the higher module (Rao et al.).

In our experiments, the subject perceived the simulated complex motions so that it departed away from the perfect complex motion induced by self-motion. So we will try to adopt the idea of top down negative feedback and explain the mechanism of our illusory effect.

Firstly, we need to assume that the perceived direction of random dot motion is given by the direction of the population vector in the MT area (fig. 6). Since the direction tuning of MT neurons have a finite width, not only the neurons with the exact preferred direction, but also the neurons with slightly different preferred directions become active when given a visual input with motion (fig. 6 (b)). Here, the preferred direction of a neuron is given by the direction of the vector and the activity is described by the length of the vector. The population vector which we assume to point the perceived direction is calculated by the vector sum of the individual vectors (fig. 6 (c)).

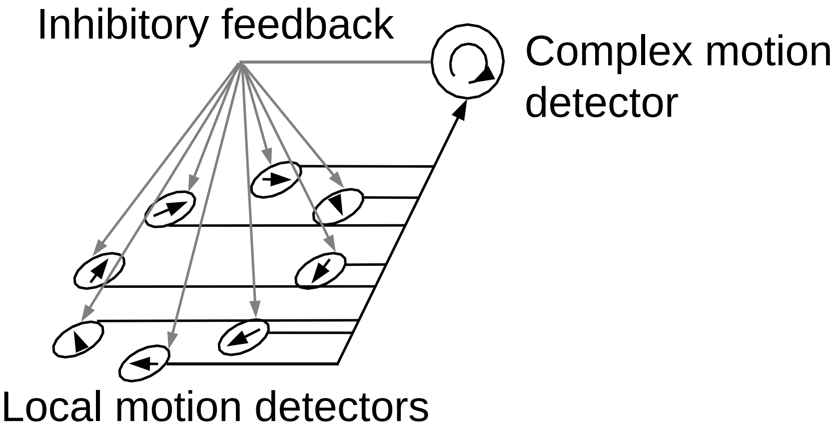

Figure 7: Schematic drawing of the possible neural mechanism which explains the novel visual illusion for rotation.

The proposed mechanism of the illusory effect is given in fig. 7 which illustrates the assumed connection among local motion detectors and a complex motion detector. We assume excitatory bottom up connections from local motion detectors to the complex motion detector, while top down connections are assumed negative. When some proportion of local motion detectors are activated by visual input, the complex motion detector become activated. Accordingly, negative feedback is sent to the local motion detectors which compose the spatial template of the activated complex motion detector. For example in the first quadrant, negative feedback is given to the directions shown in fig.8 (b). Therefore, the population vector in the left upper diagonal region is shifted downwards and population vector in the right lower diagonal region is shifted to the right, resulting in an incurve bias. Same mechanism can be adopted to explain the illusory effect for expansion and contraction.

Finally, the possible role of this inhibitory top down projections from MST to MT may be a mechanism to extract the deformation component from the optic flow field. In other words, a mechanism to separate a figure in motion and background while in self-motion. It is our future work to clarify the properties of this system.

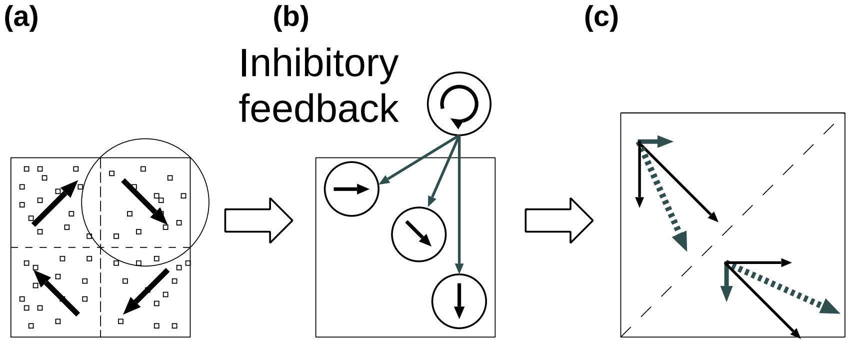

Figure 8: post synaptic potential in a two dimensional phase space of time and dendritic space

References

[1] Gatass, R. and Gross, C.G. (1981) J. Neurophysiol. 46, 621-638.

[2] Gibson, J. (1950) Houghton Mifflin, Boston.

[3] Girosi, F., Jones, M. and Poggio, T. (1995) Neural Comp. 7, 219?269.

[4] Grossberg, S. (1987) Cognitive Science 11, 23-63.

[5] Morrone, C., Blurr, C. and Vaina, M (1995) Nature 376, 507-509.

[6] Perone, A., Stone S. (1998) J. Neurosci. 18, 5958-5975.

[7] Rao, N. and Ballard H. (1999) Nature Neurosci. 1, 79-87.

[8] Saito, H. (1986) J. Neurosci. 6, 145-157.

[9] Vapnik, V.N. (1998) Statistical Learning Theory, Wiley.