

Neural Correlation of “Global-first” Topological Perception: Anterior Temporal Lobe

Tiangang Zhou · Jun Zhang · Lin Chen

Received: 8 April 2008 / Accepted: 1 September 2008
© Springer Science + Business Media, LLC 2008

Abstract The “global-first” topological approach considers that long-range apparent motion (AM) works by abstracting global form invariants, and hence is actually associating with form perception rather than motion perception. This hypothesis has been verified by the behavioral finding of motion affinity between topologically equivalent figures as well as the fMRI finding of neural correlation of AM at the anterior temporal lobe (ATL). Such interaction of behavior and brain imaging continues to impact on efforts towards a theory on perceptual objects: the topological definition of objects and the ATL neural correlation of the formation of object representation.

Keywords Global-first · Topological properties · Apparent motion · Perceptual object · fMRI · Anterior temporal lobe (ATL)

AM has received considerable attention since the beginning of the 20th century, when Gestalt psychology started. This illusion of motion, like real motion, is a common phenomenon in everyday life, as found in movies, television, video displays, and neon advertisements. The phenomenon of AM is compelling and universal: for instance, two “motionless” figures with arbitrary shapes may be separated in a distance of more than 10° of visual angle. Yet their alternation at a proper (actually a large range of) rate results in clear and continuous

illusion of motion, even though there was no corresponding physical motion in the two figures. Why does our visual system demand such illusion or what, if any, is the ecological function of long-range AM? Despite its long history and obvious theoretical and practical interest, a general-purpose theory of long-range AM that can account for its ecological functions is lacking.

In the past three decades, to address the fundamental question of “where visual processing begins” (Pomerantz 1981) or “what are the primitives of visual representation” (Chen 1982), we have developed a “global-first” (term coined by Wolfe 2001) topological approach to perceptual organization (e.g., see Chen 2005, for a review). Most traditional models of vision are “local-first”: detecting local features (such as oriented line segments) first and then integrating them, typically using attention, to build objects. In contrast to the local-first way of thinking, the global-first approach claims that topological invariants, which constitute a formal description of global, Gestalt-like operations, are the most primitive ones and are extracted at the very beginning of visual processing. A fairly large set of experimental results collected within a variety of paradigms and measures have converged at a global-first functional hierarchy, in which visual processing is *global-to-local*: the more global (or stable) a form invariant is the earlier its perception occurs, with topological perception being the most global and occurring the earliest (e.g., see Chen 2005, for a review).

This framework of the topological structure and functional hierarchy has highlighted a series of novel empirical predictions about many long-standing issues related to perceptual organization. One of the predictions is a topological account for long-range AM, namely that long-range AM works by extracting global form invariants, and hence is essentially associated with form perception rather than motion perception (Chen 1985).

T. Zhou (✉) · L. Chen
State Key Laboratory of Brain and Cognitive Science,
Institute of Biophysics, Chinese Academy of Science,
Datun Road #15, Chaoyang District,
Beijing 100101, PR China
e-mail: tgzhou@cogsci.ibp.ac.cn

J. Zhang
Department of Psychology, University of Michigan,
Ann Arbor, MI 48109, USA

The last two decades has witnessed a long journal for this topological account to be tested. That long-range AM is actually associated with form perception, as argued by topological account, is counter-intuitive, and appears to conflict with common knowledge about visual motion and the biological organization of the visual system. Thus, our recent fMRI finding is a welcome breakthrough that provided biological evidence for this topological account, beyond behavioral evidence (Zhuo et al. 2003). For the purpose of the special issue of the new journal of *Brain Imaging and Behavior*, the present paper will review how our efforts for understanding long-range AM essentially benefited from such long-standing and productive interaction between brain imaging and behavior, and how such interaction continues to impact on the future study of long-range AM in particular, and of the conceptualization of perceptual object in general.

The topological analysis of long-range AM

The correspondence problem

It has been commonly accepted that at the core of understanding AM lies the correspondence problem. That is, in the process of perceiving AM, one has to establish, at some level, a correspondence in which the visual system somehow identifies a shape in one display with its match in a later display, even though the shape may not be identical in the two displays. The starting question in the study of AM therefore is: what are the constituents of a stimulus that are matched by the correspondence process? For example, following the local-first approach, a well-known theory of AM (Ullman 1979) concluded that while some simple components, such as edge and line segments, are taken to be correspondence tokens matched in producing AM, there is no indication that structural forms are part of the correspondence tokens. Such local-first account continues to influence research in motion perception.

Invariants over shape-changing transformations

One distinguishing aspect of AM is that when one perceives AM, one perceives not only translation and rotation of rigid shapes but also intriguing “plastic deformations” occurring when AM is produced by dissimilar pairs. Under the condition of shape-changing transformations, it is difficult to imagine how simple features, such as line segments, act as correspondence tokens, because line segments making up a figure lose their identity as the figure changes, and hence lose their qualification for being correspondence tokens.

The phenomenon of shape-changing transformations further raises a key question in understanding of the correspon-

dence problem: What kinds of invariants survive such shape-changing transformations, and are consequently used by the visual system to determine that two figures represent the same object in a correspondence match no matter how different they are in shapes?

What is needed is a proper formal definition of shape-changing transformations and invariants over shape-changing transformations, beyond intuitive approach. Until the intuitive notions of these Gestalt-inspired concepts become properly and precisely defined, proposed principles of formulating correspondence tokens cannot be entirely testable.

Topology and its applicability to correspondence problem

Topology is a branch of mathematics that aims at studying invariant properties and relationships over continuous and one-to-one transformations, termed topological transformations. Intuitively, topological transformations can be imagined as arbitrary “rubber-sheet” distortions, in which neither breaks nor fusions can happen however changed in shape the “rubber-sheet” may be. From the perspective of topological analysis, solid figures (e.g., a solid triangle, a disk, and an S-like figure), although phenomenally different in appearance, are equivalent to each other, since any one of them can be modified to match any other by performing “rubber-sheet” distortions. On the other hand, “rubber-sheet” distortions cannot create or destroy holes. Thus, a ring containing a hole and a disk containing none are topologically different.

From the perspective of topological approach, shape-changing transformations observed in the phenomenal world of AM may be formally described as “rubber-sheet” distortions, i.e., the topological transformation. Invariant attributes of a shape projected onto retina over shape-changing transformations may be described as three kinds of topological properties in two-dimensional manifolds: connectivity, the number of holes, and the inside/outside relationship. If the percept of shape-changing AM indeed represents an expression of topological perception, one might expect topological properties, such as holes and the inside/outside relationship, to be good candidates for correspondence tokens in shape-changing AM (Chen 1985)

Behavioral evidence for topological account for ecological function of long-range AM

As discussed above, the topological approach suggested that topological properties may serve as correspondence tokens in producing AM. A competing motion technique was adapted to test this topological structure in AM (Chen 1985). Two stimulus displays were successively presented.

The first contains a single figure in the middle, and the second, two figures located on either side of the middle. For each presentation the subjects were asked to make a forced choice: motion from the middle figure to the figure on the right or to the figure on the left.

A series of stimulus displays (Fig. 1) were designed to manipulate topological variations, and to control for local features commonly considered in the study of vision. For one example, in (a) the triangle is topologically different from the arrow and stimulus C in holes, and in (b) the stimulus B is topologically different from stimulus A and C in holes and connectivity. However, the three figures in (a) or (b) were all made up of exactly the same three or five line-segments. Thus, line-segments as well as all local features based on these line-segments, such as luminous flux, and spatial frequency components, were well controlled. In addition, (b) was particularly designed to control for terminators: stimuli A and B had the same number (two) of “terminators” but differed in holes, and stimuli A and B differed in terminators (two vs. three) but were topologically equivalent.

As shown in Table 1, subjects displayed a strong preference for motion from a figure to a topologically equivalent figure. In the series of experiments, this motion affinity between topologically equivalent figures was well established by converging evidence. It was generalized to different kinds of topological properties, including the number of holes, connectivity, the inside/outside relation. Also, the holes presented in the stimuli differ in shapes [triangular hole in (a), circular hole in (c) and (e), and square hole in (b) and (d)], indicating that the holes perceived as an abstract topological entity contributed to the motion preference. Furthermore, to demonstrate the topological specificity of the motion affinity between topologically equivalent figures, this topological preference was tested systematically against a broad spectrum of local geometric properties, involved orientation of edge, distance, luminous flux, mirror-symmetry, parallelism, and collinearity. A major challenge to the study of the topological perception is that there can be no two figures that differ only in topological properties, without any differences in local features. Thus, one cannot test for the role of

Fig. 1 Illustration of the stimuli for apparent motion in (Chen 1985)

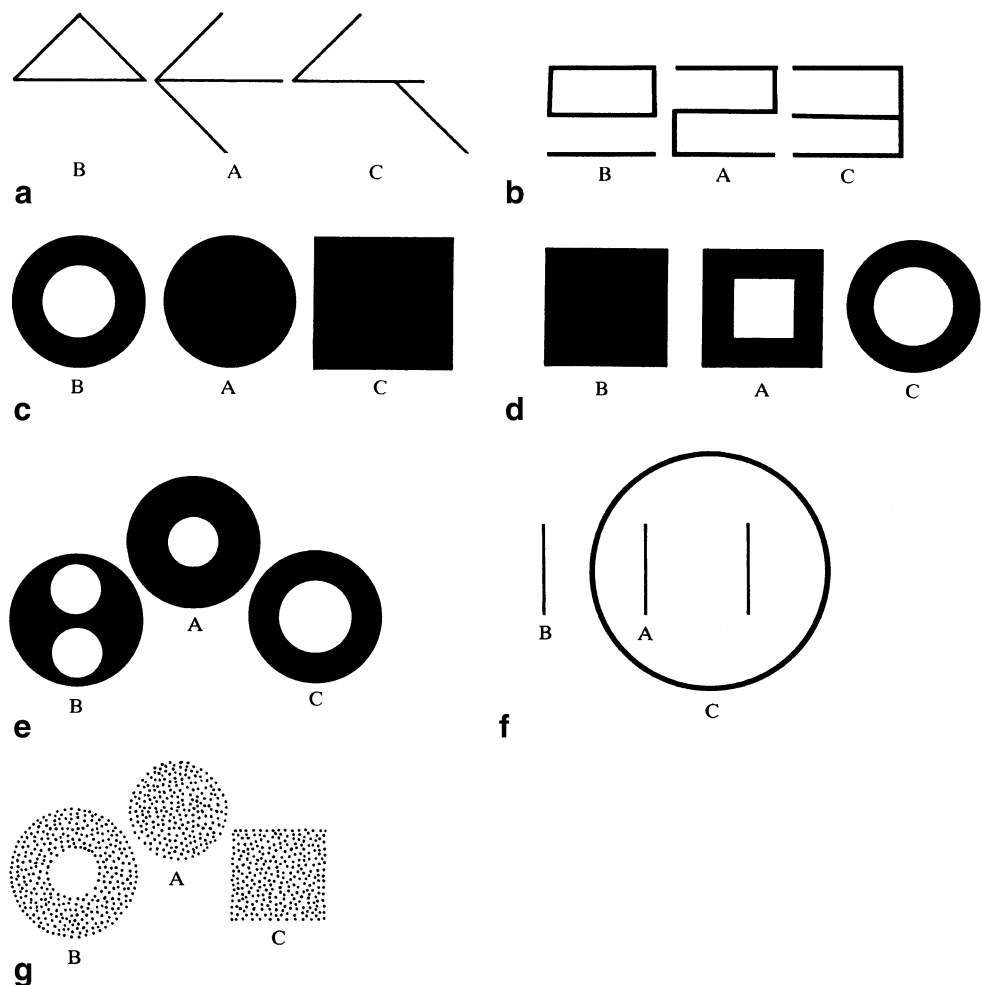


Table 1 Percentages of reports of motion from each middle figure to a figure which has the same topological invariants

Item pair	Subjects				Mean
	1	2	3	4	
a	92	83	75	75	81
b	83	92	83	83	85
c	83	92	92	83	88
d	92	83	92	92	90
e	83	100	83	92	90
f	92	92	83	83	88
g	83	92	83	75	83

The differences in average percentages between topologically equivalent and not equivalent figures are all significant, $p < 0.001$.

topological properties in AM in complete isolation. The problem was minimized through the careful design of the stimuli in a series of experiments to prevent subjects from using non-topological properties, including line segments, spatial frequency components, terminators, angles, intersections, and perimeter length, to perform the topological tasks. These nontopological features cannot, therefore, explain consistently the motion affinity between topologically equivalent figures. The topological account is the only one that explains, in a unified manner across all stimuli used, the topological preference for AM.

In summary, the topological analysis of the correspondence problem in shape-changing AM and the behavioral finding of the motion affinity between topologically equivalent figures led us to hypothesize that, with respect to its ecological function, long-range AM works by abstracting global form invariants, and hence is actually associating with form perception rather than motion perception (Chen 1982, 1985).

Predictions on neural correlation of long-range AM

Even though the behavioral experiments that revealed the topological preference for motion were convincing and well controlled, the conclusion that long-range AM is actually associated with form perception rather than motion perception still seems to be incompatible with the current dominant view that AM and real motion are detected by the same visual channels. In the two visual pathways of cortex (Ungerleider and Mishkin 1982), which are also often called the “what” or ventral pathway and the “where” or dorsal pathway, the ventral visual pathway is commonly considered to be involved in form and object perception, while the dorsal visual pathway, in motion and spatial perception. Specifically, it is commonly accepted that area MT + /V5, an area important for motion perception but not

form vision, is activated by both real motion and various illusions of motion.

According to the hypothesis of the two visual pathways, if the conclusion that long-range AM works by extracting form invariants and hence is associating with form perception is true, two predictions regarding neural correlation of long-range AM follow: (1) If long-range AM would actually be form perception rather than motion perception, then long-range AM should activate the ventral form pathway but not MT + /V5 area; (2) If long-range AM would essentially work with extracting form invariants, activation at the ventral form pathway should change along with changes of structural stability of the form invariants.

This prediction is essentially based on the topological account for AM. However, as emphasized before, this prediction appears to conflict with the current dominant view, which considers AM and real motion to be detected by the same visual channels, form perception to be of minor importance in determining AM, and motion-sensitive areas MT + /V5 in the “where” pathway to be activated by both real motion and various illusions of motion. This prediction provided, therefore, a stimulating challenge to this topological account. Nevertheless, despite its obvious importance for understanding long-range AM, the prediction was tested only recently about two decades after the prediction was made. It is not easy, if not impossible, to test this prediction with behavioral approaches alone. On the other hand, this prediction provides a good theory-driven question to guide the fMRI method with what to look at and where to look.

Neural correlation of long-range AM: ATL

Recently we used fMRI to investigate human cortical areas mediating long-range AM. The results indeed confirmed the above prediction based on the topological account for AM. We found that long-range AM generally and primarily activated the ATL, a late destination of the visual form pathway, but not area MT + , and this fMRI activation varied along with changes of form stability, in a manner similar to the hierarchy of geometries stratified with respect to their structural stability (Zhuo et al. 2003). In this paper, while we will extend reviewing the data reported in (Zhuo et al. 2003) in general, we will highlight an analysis of the relation between area MT + and long-range AM, namely that long-range AM activates only the ATL rather than area MT + , to argue for the unique importance of the ATL in the neural correlation of long-range AM.

Long-range AM vs. flicker In one of these experiments, the activation stimulus was two squares separated by about 10° and presented in alternation so as to produce AM; and the

baseline stimulus was the same two squares but presented simultaneously so that no AM but only flicker was generated [Fig. 1a in Zhuo et al. 2003]. The fMRI result showed activation in, in addition to the lateral occipitotemporal cortex, the anterior temporal gyri. This result was a surprise because the ATL is a late destination of the visual form pathway, which is anatomically far removed from area MT + and the where pathway. However, these results left the two perceptual components of long-range AM and flicker for further dissociation, because the stimuli of two alternating squares not only generated the percept of AM but also involved physical flicker. These findings, therefore, raised a further issue: did long-range AM *per se* or the confounding flicker produce the activation in area MT +?

Kinetic forms, controlling for flicker To address this issue, kinetic forms were used in producing AM. Kinetic shapes are defined solely by spatiotemporal correlations rather than luminance differences. Thus all luminance-based features were well controlled. With respect to kinetic forms, another characteristic is worth emphasizing. Although AM can be produced by kinetic shapes defined by static random dots against a background of dynamic noise, no AM can be produced at all by shapes defined by dynamic dots against a background of static random dots (Fig. 2). In the experiments

using luminance-contrast forms, two figures must be presented alternately in order to produce AM. To use the same two figures as a baseline condition that controls for possible confounds (for example, luminance flux) but without invoking any AM, one normally has to present the two figures simultaneously. Therefore, by limiting the stimuli to luminance-contrast forms, it would be difficult to equate both spatial and temporal aspects of stimulus presentation between the two conditions.

We made use of this asymmetry that prevails in kinetic forms in AM, to provide some additional controls for presentation conditions. The activation task of two kinetic squares was compared with the baseline task of the same two kinetic squares except for a reversal of the relation of figure and surround. Because the kinetic forms in the baseline condition produced no AM at all, the task and baseline stimuli could be presented with identical presentation conditions. If it is true that long-range AM *per se* activates the anterior temporal gyri in general and without involving area MT +, two predictions follow: (1) Like the luminance-contrast forms, the kinetic forms would also produce activation at the ATL; but (2) unlike the luminance-contrast forms, the kinetic forms would not produce activation in area MT +. The fMRI results indeed confirmed these two predictions: Comparing the AM and

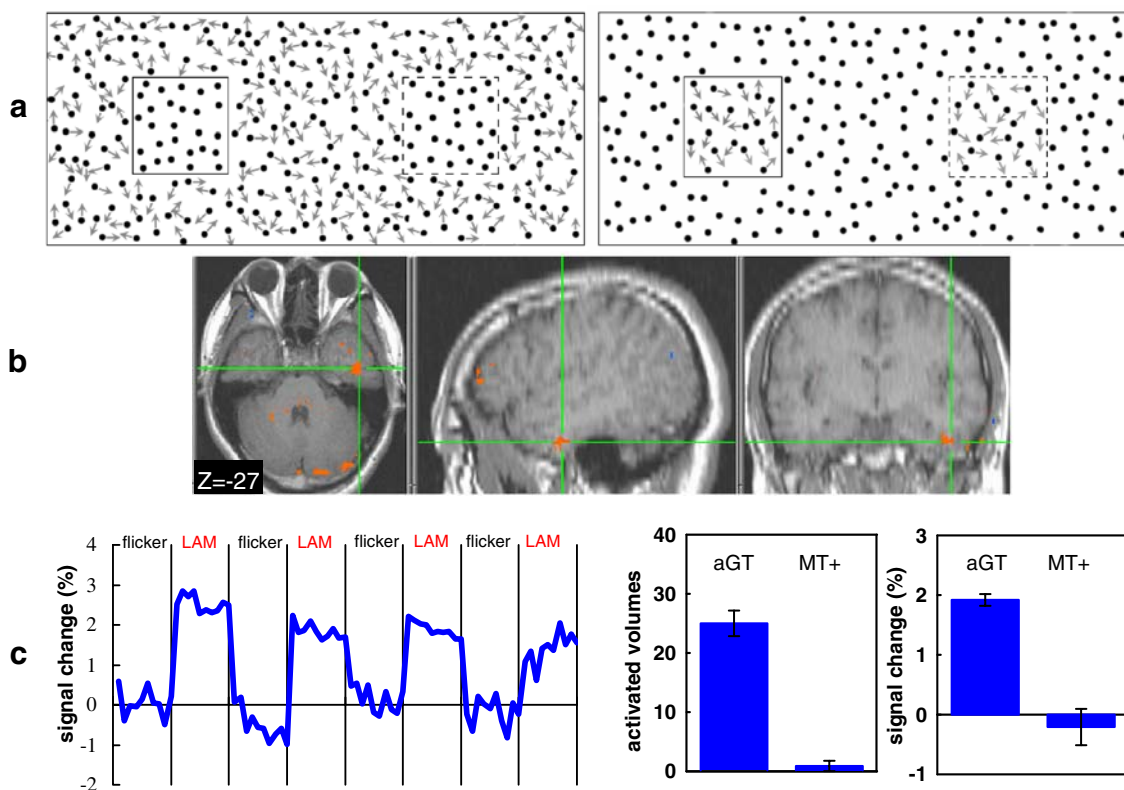


Fig. 2 (a) Illustration of the kinetic stimuli in (Zhuo et al. 2003). Left of A shows kinetic shapes defined by static random dots against a background of dynamic noise, and right of A shows shapes defined by

dynamic dots against a background of static random dots. (b) & (c) were their results

no AM conditions, significant activation was still found in the anterior temporal gyri, and no significant difference in activation in area MT + was found.

S-like figure vs. ring, controlling for local features For further testing whether the activation in the ATL was specific to long-range AM, two pairs of figures, an S-like figure vs. a disk and an S-like figure vs. a ring [Fig. 3a, in Zhuo et al. 2003] were compared with the baseline task of two identical S-like figures, in producing long-range AM. The S-like figure is topologically different from the ring in holes. However, the S-like figure was made to approximate the area of the ring. Even though the local-feature differences between the S-like figure and the ring (such as luminous flux difference, spatial frequency components, and perimeter length) were minimized, this topologically different pair still caused stronger activation in the ATL than the topologically equivalent pair of S-like figure and disk.

Activation correlated with form stability As predicted before, if the activation of the ATL indeed occurred because long-range AM is associated with form perception, then varying the form properties should influence the activation pattern. Five pairs of figures (including the base-line pair A) were designed to produce AM (Fig. 3). The differences between the two figures in pairs B to E represent different levels of form stability. In ascending order from pairs B to E, they differ in Euclidean geometry (a square vs. a parallelogram), affine geometry (a square vs. a trapezoid), projective geometry (a square vs. a solid circle), and finally topology (a square vs. a ring) with the highest stability. The fMRI data shows that the activated cortical volumes as well as the amplitudes of signal changes in the ATL increased monotonically with increasing levels of the stability of structural differences in the forms (Fig. 3b through d). This result suggests that the more stable in terms of structural difference between two forms, the greater the magnitude of cortical activation required to produce the perception of AM between the two forms. Specifically, pairwise comparisons indicated that Pair E, representing the highest stability (topological difference), caused the strongest activation in comparison with Pairs D, C, and B.

Towards the topological definition of perceptual objects and the ATL neural correlation of the formation of object representation

The correspondence problem in shape-changing AM is essentially, we believe, related to a more general notion of perceptual *objects*. What is a perceptual object? This

question seems to be straightforward yet its answer has become one of the most central and also controversial issues in many areas of cognitive sciences. However, after decades of efforts, we do not yet know exactly what counts as an *object* in the first place.

The concept of perceptual objects and correspondence problem

To appreciate the strength of the notion of perceptual objects, we may appeal to the analysis of the phenomenon of shape-changing AM. Two dissimilar shapes displaced intermittently across the visual field (for example, a triangle moves and changes its shape simultaneously to become a disk or vice versa) will nonetheless give an illusion of continuous motion. As emphasized above, the existence of AM under such shape-changing conditions implies a correspondence match, in which the visual system somehow identifies which part or attribute of stimuli presented successively in different frames represents the same object. In this regard, the correspondence problem is essentially related to the concept of perceptual objects. In the more general sense of invariance perception, long-range AM may represent an expression of the formation of object representation. This analysis of the relationship between correspondence problem and the concept of objects illustrates the dynamic nature of object representations, namely that an object appears to be something that preserves its identity over shape-changing motion, even if the notion of an object appears to be a static one. To understand the concept of perceptual objects turns out to face the same key question to be answered as the correspondence problem: what kinds of invariants of a shape survive such shape-changing transformations, and consequently are used by the visual system to determine the shape, despite the fact that it is subjected to shape-changing motions, nevertheless representing the same object?

As a matter of fact, shape-changing transformation observed in long-range AM is often observed in everyday life. An object in the natural environment, say a bird flying, is often subjected to shape changes due to its motion (which may be non-rigid) or to changes of illumination. Phenomenal impression will be that regardless of any changes in its featural properties (e.g., location, orientation, size, and shape), the bird retains its identity as an object over such deformations. Generally, shape-changing motion introduces a fundamental aspect in visual perception, that is, in Marr's (1982) term, "the consistency of an object's identity through time". For rigid bodies, the motion correspondence process in time domain and the stereo correspondence process in space domain, are essentially equivalent to each other. Thus, a computational approach, applicable to the correspondence problem in stereopsis, seems to be also suitable

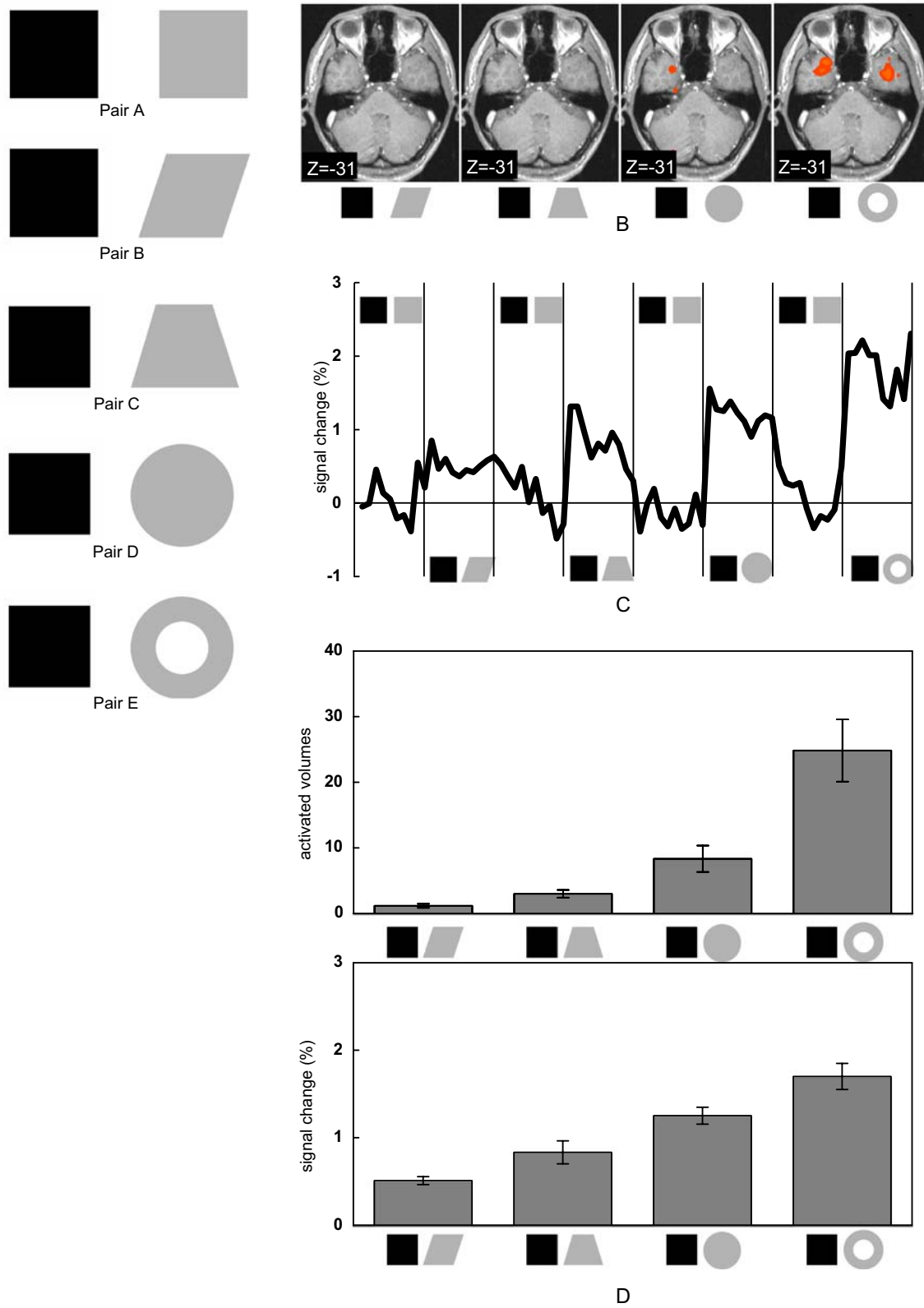


Fig. 3 Stimuli and results of functional hierarchy experiment in (Zhuo et al. 2003)

for the motion correspondence process. However, as Marr (1982) realized, for shape-changing bodies, a new theory may be necessary for motion correspondence process, because, in shape-changing motion, time introduces a new fundamental issue. This issue is not part of the structure-from-motion problem because the precise details of an object's structure are not relevant. Nevertheless, despite the object identity problem being at the core of understanding the correspondence problem, Marr did not provide any hint as to what this new theory might be.

It is the global-first topological approach that provides a new theory for the correspondence problem in shape-changing motion as well as for the definition of perceptual objects.

The topological definition of perceptual objects

As emphasized in the topological analysis of AM, from the perspective of the topological approach, the phenomenal impression of shape-changing transformations, which intuitively can be imagined as arbitrary “rubber-sheet” distortions, may be formally described in terms of topological transformations. That is, a perceptual object can be defined as something that keeps its topological structure over time. The topological approach, therefore, ties the formal definition of perceptual objects to invariance over topological transformation, and the core intuitive notion of a perceptual object—the holistic identity preserved over shape-changing transformations—may be precisely characterized as topological invariants, such as connected components. For example, as we speak of “an object” in a picture, we usually imply that it is connected.

The topological definition of perceptual object, inspired by the study of long-range AM, provides a framework of global-first topological approach to a fundamental and general question that can be asked about any cognitive process: what are the primitive units over which the cognition process operates? The topological definition therefore highlights its broad applicability to various issues at different levels of cognition, much more beyond AM. For example, with respect to the object-based theory of selective attention, the topological definition of objects predicts that topological properties are basic constraints in forming perceptual objects that attention selects. In particular, the topological change, such as in holes, will be perceived as the emergence of a new object that will capture attention and impact on attention-related performance. This prediction has been verified in the studies carried on in our Laboratory of, for example, multiple object tracking, pre-cuing attention, and attention capture (Zhou et al. 2004; Zhou and Chen 2008); binocular rivalry (Chen et al. 2008); and attention blink (Qian et al. 2007).

Neural correlation of topological perception: ATL The finding of the involvement of the ATL in long-range AM has stimulated a great enthusiasm to explore the role of the ATL in the formation of object representation defined by topological constraints. The fMRI method has been widely applied to measure cortical activation involved in topological perception and object-based cognitive processing under various conditions. These fMRI studies have involved, for example, in addition to long-range AM (Zhuo et al. 2003), hemispheric asymmetry (Wang et al. 2007), multiple object tracking (Zhou et al. 2004; Zhou and Chen 2008), visual search (Zhou et al. 2006), and binocular rivalry (Chen et al. 2008). It is interesting to see that all these fMRI experiments demonstrated that the ATL is involved in global topological perception and the formation of object representation.

In summary, on the one hand, the global-first topological approach highlights a series of novel behavioral findings at different levels of cognition, converging at that the extraction of topological properties serves as the starting point of the formation of an object representation. On the other hand, our brain imaging studies have revealed the involvement of the ATL, a late destination of the visual form pathway, in global topological perception. This contrast of global-first in behavior and late destination in neuroanatomy is likely to raise far-reaching issues regarding the formation of object representations in particular, and the fundamental question of “where to begin” in general.

Stimulated by such progresses made in our Lab, a scientific collaboration movement is extending to institutes of the Chinese Academy of Sciences, universities, and hospitals, in China, as well to Hong Kong, and US. The collaboration is motivated by a common scientific goal of targeting mainly the ATL, and aims at promoting the development of methodology and technology for integrating fMRI, MEG, ERP, and TMS plus human recording, and monkey fMRI.

Acknowledgments This research was supported by Ministry of Science and Technology of China Grants 2005CB522800 and 2004CB318101, National Nature Science Foundation of China Grant 30621004, and Knowledge Innovation Projects of Chinese Academy of Sciences (L.C.).

References

- Chen, L. (1982). Topological structure in visual perception. *Science*, 218, 699–700. doi:10.1126/science.7134969.
- Chen, L. (1985). Topological structure in the perception of apparent motion. *Perception*, 14, 197–208. doi:10.1068/p140197.
- Chen, L. (2005). Topological approach to perceptual organization, lead paper in *Visual Cognition*, 12, 553–637.

- Chen, Y. Q., Zhou, T. G., & Chen, L. (2008). The sensitivity of binocular rivalry to the topological change (in preparation).
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W. H. Freeman.
- Pomerantz, J. R. B. (1981). Are complex visual features derived from simple ones? In E. L. J. Leeuwenberg, & H. F. J. M. Buffart (Eds.), *Formal theories of visual perception*. Chichester: Wiley.
- Qian, W. L., Zhou, K., & Chen, L. (2007). The emergence of new object produces attention blink: A topological account. In *Proceedings of the Second Annual Symposium of Brain and Cognitive Science, China*.
- Ullman, S. (1979). *The interpretation of visual motion*. Cambridge: MIT.
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. G. Ingle, M. A. Goodale, & R. J. Q. Mansfield (Eds.), *Analysis of visual behaviour* (pp. 549–586). Cambridge: MIT.
- Wang, B., Zhou, T. G., Zhuo, Y., & Chen, L. (2007). Global topological dominance in the left hemisphere. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 21014–21019. doi:10.1073/pnas.0709664104.
- Wolfe, J. (2001). Personal communication.
- Zhuo, Y., Zhou, T. G., Rao, H. Y., Wang, J. J., Chen, M., Zhou, C., et al. (2003). Contributions of the visual form pathway to long-range apparent motion. *Science*, 299, 417–420. doi:10.1126/science.1077091.
- Zhou, K., Lu, S. N., Hu, S. Y., & Chen, L. (2004). Wholes, holes, and objects in selective attention. In *Proceedings of ICP2004* (Invited symposium: Brain mechanisms for selective attention).
- Zhou, K., & Chen, L. (2008). The topological approach to perceptual organization: Where visual processing begins. In *Proceedings of ICP2008* (Invited address No. IA-016).
- Zhou, T. G., Cai, P., & Chen, L. (2006). The primitives of visual representation: Evidence from visual search. In *Proceedings of the First Annual Symposium of Brain and Cognitive Science, China*.