

Theory of Cognitive Pattern Recognition

Youguo Pi, Wenzhi Liao, Mingyou Liu and Jianping Lu
*School of Automation Science and Engineering, South China University of Technology
 Guangzhou, Guangdong,
 China*

1. Basis of cognitive psychology related to pattern recognition

1.1 Perception and its constancy

Born and developed in the middle of 1970's, cognitive science is a kind of intersectional and integrative science aiming to study both the working principle and the developing mechanism of human brain and psyche. It is a product from the processes of intersection, infiltration and aggregation of such sciences as psychology, computer science, neurology, linguistics, anthropology, philosophy, and so on.

As one of the important parts of cognitive science, cognitive psychology[1-6], developed in the middle of 1950's, is a kind of psychology making the view of information processing as the core, thus also named information processing psychology, and a kind of science studying the processes of transforming, processing, storing, recovering, extracting and using information through sense.

Perception has always been an important studying field of psychology. Cognitive psychology treats perception as the organization and explanation of sense information, and the process of acquiring the meanings of sense information. Correspondingly, this process is treated as a series of consecutive information processing, and the ability of the process depends on the past knowledge and experience.

We can cover a building far away by just a finger, it means that the image of finger formed on the retina is bigger than that of the building. But if we move away the finger and first look at the building then the finger, we will feel the building is much bigger than the finger anyway, that indicating a very important feature of perception-constancy. The constancy of perception refers to perception keeps constant when the condition of perception changes in a certain range [7]. In the real world, various forms of energy are changed while reaching our sense organs, even the same object reaching our sense organs. Constancy in size and shape keeps our lives normal in this daedal world. Although an object sometimes seems smaller or bigger, we can recognize it. Constancy is the basis of stable perception of human to the outside. For instance, students can always recognize their own schoolbag, no matter far away (assuming it is visible) or close, overlooking or upward viewing, or looking in the front or sides. Although the images formed in the retina under the different conditions mentioned above are different from each other student's perceptions of this object are the same schoolbag.

Constancy in size and shape are two main types of the perception constancy. Perception constancy in size means that although the size of object images shot on the retina change, human perception of the size of object keeps constant. The size of image on the human retina directly depends on the distance between the object and our eyes.

Source: Pattern Recognition Techniques, Technology and Applications, Book edited by: Peng-Yeng Yin, ISBN 978-953-7619-24-4, pp. 626, November 2008, I-Tech, Vienna, Austria

For example, a man is coming toward you from far away, but after you recognize who he is, although his image on your retina is growing bigger and bigger as he is getting closer and closer to you, your perception of the coming person has nearly no change but just that guy. This perception, of course, has boundary, the farthest boundary are where you can recognize the person. Is there any nearest boundary? Suppose a very tall man, which is double or triple of you, gets close to you, you can only see his leg, at this time you can not recognize who he is. When he returns back facing you, as the distance between you and him increases, the image you have is closer and closer to his panorama, then you can recognize him. Therefore we may interpret the size constancy of perception as this: in the condition that image information is enough to recognize the pattern, the size of the image doesn't affect human's perception.

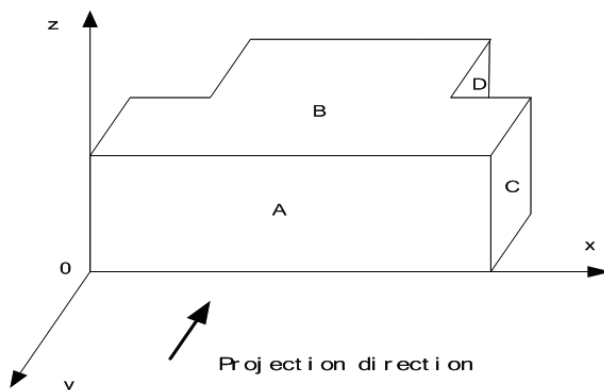


Fig. 1. The constancy of the perception

The shape constancy of perception means that in perception, although the shape of the object image shot on retina changes, people's perception of the shape of object stays constant. The shape of image on human retina directly depends on the angle of view between the object and eyes. As shown in figure 1, when the object is projected in the normal direction of plane A, we can only see plane A without the whole shape of this object. When we move the direction of view along the positive way of x and z axis, and can see plane A, B, C, and D. No matter what the size and proportion of these four planes, we can still recognize the object. This is shape constancy of perception. We now can interpret it as follows: when image information is enough to recognize the pattern, the changes of image shape don't affect human perception of the object.

1.2 Pattern recognition

Pattern recognition is one of the fundamental core problems in the field of cognitive psychology. Pattern recognition is the fundamental human cognition or intelligence, which stands heavily in various human activities. Tightly linking with such psychological processes as sense, memory, study, and thinking, pattern recognition is one of important windows through which we can get a perspective view on human psychological activities.

Human pattern recognition can be considered as a typical perception process which depends on knowledge and experience people already have. Generally, pattern recognition refers to a process of inputting stimulating (pattern) information and matching with the

information in long-term memory, then recognizing the category which the stimulation belongs to. Therefore, pattern recognition depends on people's knowledge and experience. Without involving individual's knowledge and experience, people cannot understand the meanings of the stimulating information pattern inputted, then neither possible to recognize the patterns, which means to recognize the objects. The process which a person distinguishes a pattern he perceives with others and identifies what it is means pattern recognition. Current cognitive psychology has proposed such theoretical models or hypothesis as the Theory of Template (Model of Template Matching), the Theory of Prototype (Model of Prototype Matching), the Theory of Feature (Model of Feature Analysis), and so on.

(1) The Theory of Template

As the simplest theoretical hypothesis in pattern recognition, the Theory of Template mainly considers that people store various mini copies of exterior patterns formed in the past in the long-term memory. These copies, named templates, correspond with the exterior stimulation patterns one by one. When a simulation acts on people's sense organs, the simulating information is first coded, compared and matched with pattern stored in brain, then identified as one certain pattern in brain which matches best. thus the pattern recognition effect is produced, otherwise the stimulation can not be distinguished and recognized. Because every template relates to a certain meanings and some other information, the pattern recognized then will be explained and processed in other ways. In daily life we can also find out some examples of template matching. Comparing with template, machine can recognize the seals on paychecks rapidly.

Although it can explain some human pattern recognition, the Theory of Template, meanwhile, has some obvious restrictions. According to the Theory of Template, people have to store an appropriate template before recognize a pattern. Although pre-processing course is added, these templates are still numerous, not only bringing heavy burden to memory but also leading pattern recognition less flexible and stiffer. The Theory of Template doesn't entirely explain the process of human pattern recognition, but the template and template matching cannot be entirely denied. As one aspect or link in the process of human pattern recognition, the template still works anyway. In some other models of pattern recognition, some mechanisms which are similar to template matching will also come out.

(2) The Theory of Prototype

The Theory of Prototype, also named the Theory of Prototype Matching, has the outstanding characteristic that memory is not storing templates which matches one-by-one with outside patterns but prototypes. The prototype, rather than an inside copy of a certain pattern, is considered as inside attribute of one kind of objects, which means abstractive characteristics of all individuals in one certain type or category. This theory reveals basic features of one type of objects. For instances, people know various kinds of airplanes, but a long cylinder with two wings can be the prototype of airplane. Therefore, according to the Theory of Prototype, in the process of pattern recognition, outside simulation only needs to be compared with the prototype, and the sense to objects comes from the matching between input information and prototype[5]. Once outside simulating information matches best with a certain prototype in brain, the information can be ranged in the category of that prototype and recognized. In a certain extent the template matching is covered in the Theory of Prototype, which appears more flexible and more elastic. However, this model also has

some drawbacks, only having up-down processing but no bottom-up processing, which is sometimes more important for the prototype matching in human perceptual process.

Biederman(1987,1990) proposed the theory of Recognition-By-Components, whose core assumption is that, object is constituted by some basic shapes or components, or say geometries which includes block, cylinder, sphere, arc, and wedge. Although the number of components seems not enough for us to recognize all objects, these geometries can be used to describe efficiently, for the various spatial relations of all geometries can constitute countless assembles. The Step one of Biederman's Recognition-By-Components process is extracting edges, and the Step two divides a visible object into some segments to establish the components or geometries constituting the object. The other key is that the edge information has invariant properties, based on which the components and geometries of the visible object are established.

(3) The Theory of Feature

The Theory of Feature is other theory explaining pattern perception and shape perception. According to this theory, people try to match the features of pattern with those stored in memory, rather than the entire pattern with template or prototype. This model is the most attractive one currently, the Model of Feature Analysis has been applied widely in computer pattern recognition. However, it is just a bottom-up processing model, lacking up-down processing. Therefore, it still has some drawbacks.

1.3 Memory

First、 The Description of memory

Memory is a reflection of the past experience in human brain, and, in cognitive psychology, a process of information coding, storing, and extracting in a certain condition in future. Having a big effect on human history and individual person development, memory is a gift from the nature to individual life, and also a power with which individual keeps and uses the achieved stimulating information, knowledge and experience

As a necessary condition of the intellect development, memory is the root of all intelligence. People keep past experience into their brain by memory, and then, based on experience recovering, have thinking and imagination, whose results are kept again in brain as the basis of further thinking and imagining.

Memory, in cognitive psychology, can be seen as a process of information inputting, coding, storing, and extracting, therefore, it can be separated as instantaneous memory, short-term memory, and long-term memory according to the time of storage. Recent years, more and more researchers propose to view memory as multiple memory form with different property functions formed with various forms, systems or types (Schacter 1985).

Second、 The model of memory system

In 1960's, relying on the deep research of short-term and long-term memory, researchers on cognitive psychology gradually proposed and built some memorial theory and related memorial models. Among them, the Multiple Mnemonic Model proposed by Atkinson and Shiffrin in 1968 is the most attractive one, as shown in figure 2.

In this model, memory is described by 3 kinds of memory storages: ①sensory store, limited number and very short time for the information keeping; ②short-term store, longer time of storage but still limited number to keep;③long-term store, powerful power of storage, and able to keep for a long time, or maybe even forever. However, recently cognitive psychologists usually describe these 3 kinds of storages as sensory memory, short-term

memory, and long-term memory. In this model, outside information first input into sensory registration, which has various kinds of information but probably disappears very soon. Then the information will be transferred into short-term memory, in which the information is organized by hearing, language or spoken language acknowledgement, and is stored longer than that in sensory storage. If processed meticulously, repeated, and transferring acknowledged, the information will be input into long-term memory, or else will decline or disappear.

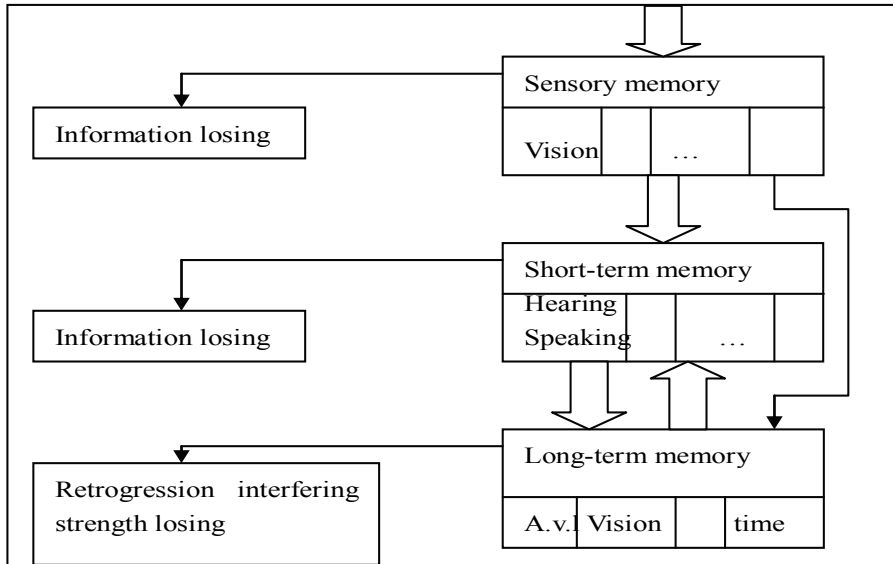


Fig. 2. The model of memory system

1.4 The expression and organization of knowledge

Human has transcendental imagination. If imagination is produced by experience and knowledge, then human’s knowledge must be organized by a certain way. Cognitive psychology describes inside knowledge attribution of individual through establishing cognitive model, which has 3 hypothetical models.

First、Hypothesis of symbol-net model

This model can comparatively indicate how every part of knowledge in human brain arrays and interacts with each other in a certain connecting mode.

In symbol-net model, conceptions are usually described as “node”, which links each other with a arrowed line, and therefore the two concepts are connected by a certain mode. In symbol-net model, we describe this relation with “up and down level”, adding with arrowed line. What needs to be attended is the arrow direction, which has some theoretical meanings in symbol-net model, as figure 3 showing.

The fundamental assumption of symbol-net model is a reflection of people’s knowledge organization, which is similar to searching among the network nodes. The search is performed one node by another along the direction of the arrows according to the form of cognitive process series, until reach the nearest node and search out the knowledge. If the

knowledge in the nearest node can answer the certain question, the search will cease, otherwise the search will continue till finding out answer or giving up.

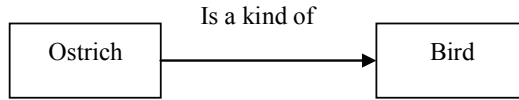


Fig. 3. The Symbol-net model

Second、 Level-semantics-net model

The Level-semantics-net Model, proposed by Collins and Quillian, is a net connecting with various elements, the node represents a concept and the arrowed line reflects the affiliation of concepts. This model indicates that every concept or node has two relationships, one is that every concept is subject to other concepts, which deciding the type of knowledge attribution and describing the affiliation with “is a kind of” relation; the other is that every concept has one or more characteristics, meaning the “have” relation of concept, as figure 4 showing.

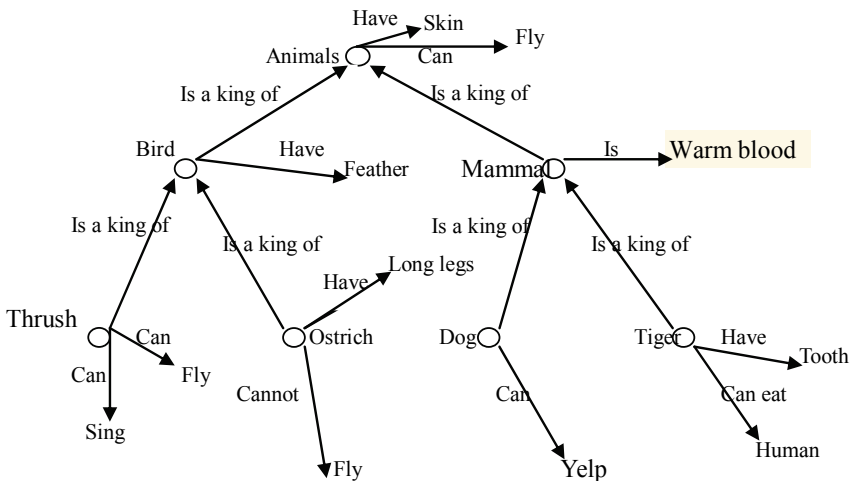


Fig. 4. The Level-semantics-net model

According to this model, the organized knowledge attribution is a level dendriform view, in which lines link nodes denoting concepts of each grade, actually in a certain extent has some imagining function. In this model, because concepts form a net according to “up-and-down” grades, every concept and characteristic locates in a specific position in the network, and the meaning of a concept depends on connecting links. According to the cognitive economic principle, the Level-semantics-net model maximizes the effective storage capability while minimizes the redundancy.

Third、 The activation-diffusion model

The core of Level-semantics-net model is the network established by the logical relations of noun concepts. This features the model clean and clear, but also causes some problems, which mainly appears that the model explains human knowledge organization and attribution assuming on logics rather than psychology. Therefore, Collins and Loftus modified the original model and proposed a new one, which is the activation-diffusion model. Giving up the level structure of the concepts,, the new model organizes concepts by the connection or similarity of semantics.

In activation-diffusion model, the knowledge stored in individual's knowledge structure is a big network of concepts, between which certain connection is established, namely some knowledge is contained in advance. Therefore, activation-diffusion model is also a kind of pre-storing model, as shown in figure 5.

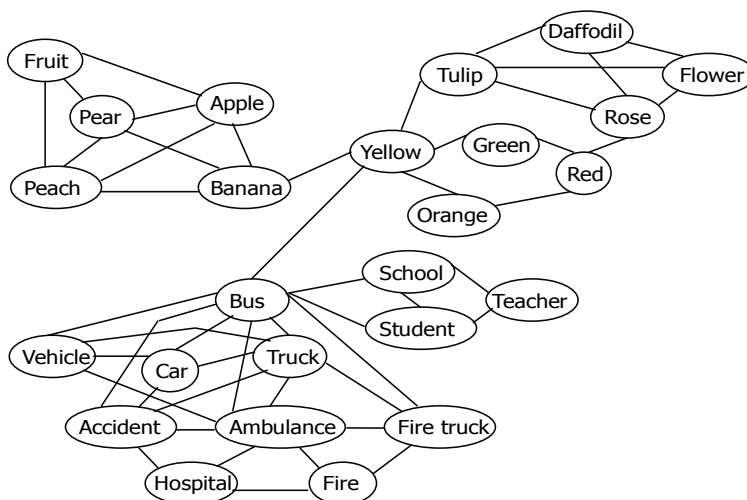


Fig. 5. The activation-diffusion model

The activation-diffusion model has two assumptions related to knowledge structure: first, the line between concepts reveals their relation, the shorter the line, the tighter their relation, and the more similar their features, for instance, "car" having tight relation with "truck", rather with "teacher", second, the intension of concept of the model is decided by other related concepts, especially the tight ones, but the features of concept is unnecessary to be stored in different grades.

1.5 The theory of topological vision

Lin Chen, involving topology into visual perception study, proposed The theory of topological vision [7]. The topology study of perceptual organization is based on a core idea and composed by two aspects. The core idea is that, perceptual organization should be interpreted in the angle of transformation and its invariance perception. One aspect emphasizes the topological structure in shape perception, which means that the global characteristic of perceptual organization can be described by topological invariance. The other aspect further emphasizes the early topological characteristic perception, which means

that, topological characteristic perception priors to the partial characteristic perception. The word "prior" has two rigid meanings: the entire organization decided by topological characteristics are basis of the perception of partial geometric characters, and topological characteristics perception of physical connectivity is ahead of perception of partly geometric characteristics.

2. Brief commentary of machine pattern recognition

Machine pattern recognition developed rapidly in the beginning of 1960's and became a new science, then has been in rapid development and successfully applied in weather forecasting, satellite aerograph explanation, industrial products measurement, character recognition, voice recognition, fingerprint recognition, medical image analysis and so on.

By now Machine pattern recognition (pattern recognition for short) mainly has two basic methods: statistics pattern recognition and structure (syntax) pattern recognition. Structure pattern recognition, based on image features of structure, accomplishes pattern recognition by using dendriform information of the layered structure of pattern and subschema. Statistics pattern recognition, which has wider application, is based on the type probability density function of samples in feature space and separates pattern statistics into types, which means pattern recognition integrated with Bayesian decision in proportion statistics, is also called decision theory recognition method.

In statistics pattern recognition, some knowledge and experience can decide the principle of classification, which means the rules of judgment. According to appropriate rules of judgment, we can separate the samples of feature space into different types and thus change feature space to type space. We separate feature space into type space while we classify the patterns. Statistics pattern recognition is based on the type probability density function of samples in feature space, and the rule of judgment of multiple statistics pattern recognition is Bayesian decision theory, aiming to minimize the expected risk of prior probability and lost function. Because nonlinear classification can be transferred into linear classification, the fact is searching the hyper plane of optimal decision. Although Bayesian decision rules solve the problem of engineering the optimal classifier, the implement has to be first settled with the more difficult problem of probability density distribution, thus research develops surrounding decision rules and probability density distribution. For the former, Rueda L G and Oommen B J's researches in recent years indicate that the normal distribution and other criteria functions with the covariance matrix unequal are linear and their classifiers is optimal[9]; Liu J N K, Li B N L, and Dillon T S improved Bayesian classifier with genetic algorithm when choosing input feature subset in classification problem[10]; Ferland G and Yeap T, studying the math structure of RTANN method, identified the condition of achieving optimal Bayesian classification with such method[11]. For the issue of probability density distribution, usual assuming density is a model with parameters like multiple normal distribution, while the parameters are estimated by the training sample. When the sample is not enough, the estimated error which is contained by distribution function will affect the precision of recognition. In order to improve the precision of recognition, Ujje H et al transformed the reference data closer to normal distribution, no matter what the distribution of original data, and found the optimal transformation in theory [12]. The emergence of statistic learning and supporting vector machine bring theoretical and methodological supplement for the transformation. Core function which satisfy the Mercer condition realizes the design of the nonlinear classifier without knowing the specific form of

nonlinear transformation[13]. Fisher judgment and principal component analysis are traditional linear methods which widely applied in pattern classification and feature extraction. The Fisher judgment [14-15]and principal component analysis[16] in recent years both based on the core function are their linear widespread. One-dimensional parameter search and recursion Fisher method can get better training result than normal Fisher judgment. Using Mercer core, we can generalize these two methods into nonlinear decision plane[17]. There are also some reports of improving the function of classifier by declining pattern overlapping with fuzzy cluster analysis[18].

Therefore, there are two main problems need to be solved in pattern recognition:

1. Because of the requirement of sample amount, statistics pattern recognition cannot function well in small sample recognition.
2. so far, the pattern recognition is mainly based on the classification mechanism of the recognized objects, rather than on the perception mechanism. In "recognition", namely in the aspect of acknowledge of objects (study), there is large difference between human perception process and limited learning ability.

3. Theory of cognitive pattern recognition

3.1 Perceptive constancy and topological invariance

In the first chapter, we generally express perceptive constancy as: in the condition that the image information of the object is sufficient to determine its pattern, the geometry changing in the size and shape does not affect people's perception for the object.

The above questions refer to a special kind of geometric properties of geometry, which involve the property of the geometric overall structure, named the topological property. Obviously, these topological properties are not related to such aspects of the geometry as the size, the shape and the straight or curved of lines and so on, which means that they can not be dealt with by ordinary geometric methods. Topology is to study the invariable property when geometry makes elastic deformation, the same as the perceptive constancy that changing in size and shape of the geometry do not affect people's perception for the object.

Now let's make a further analysis to the topological property. As mentioned above, topology embodies the overall structure of geometric features, that any changes in shapes (such as squeezing, stretching or distorting, etc), as long as the geometry is neither torn nor adhered, will not destroy its overall structure and the topological properties remain the same. The above deformations are called the topological transformation, so the topological invariability is the property keeping the same when the geometry transforms topologically. The topological property can be accurately described by the set and mapping language. The changing of the geometry M to M' (both can be regarded as a set of topological feature points) is a one-to-one mapping (therefore the overlap phenomenon will not appear, moreover new points will not be created) $f : M \rightarrow M'$, where f is continuous (that means no conglutination). Generally speaking, if both f and f^{-1} are continuous, the one-to-one mapping f , which changes M to M' , can be regarded as a topological transformation from M to M' , also f and f^{-1} are the mapping of homeomorphism. Therefore, topological property is common in the homeomorphous geometries. The geometries of homeomorphism have no differentiation in topology because their topological properties are the same.

3.2 Perceptive constancy and pattern invariance

From above discussion, we can regard the perceptive constancy as the topological invariance. As the size constancy, changing the size of geometry is actually compressing and expanding the geometry during which the topological properties of the geometry do not change. And the shape constancy means to carry unequal proportional compression and expansion on geometries. As shown in figure 1, when we make projection on the normal direction of plane A, which creates conglutination between plane A and plane B, C, D, geometric topology has been changed, so the object can not be perceived. When the projection points move along the x-axis and z-axis, we can observe that conglutination has not been created among plane A, B, C, D, the topological structure has not been changed, so the object can be perceived.

Furthermore, we will discuss perceptive constancy by using the theory of topology.

First, size constancy :

As mentioned above, as the distance between human eyes and the object changes the image sizes of geometry on the retina change, but in our minds we perceive the images of different sizes as an object, we call this kind of information processing size constancy. The explanation of size constancy is shown in figure 7. In the figure, as the distance between the eyes and the object changes, the images are named a, b, c and d respectively. In image a, the distance between the eyes and object is so near that the image of the object cannot be seen entirely, thus unable to be recognized. As the distance between the eyes and object becomes farther and farther, the images of the object on the retina become smaller and smaller, as shown in figure b, c and d. In the figure d, the distance to the eyes is too far and the image of the object is too small to recognize.

Sequence images of the object are generated on the retina as the distance between the eyes and an object X changes. Now suppose Y is the image generated on the retina at a certain distance from eyes within the human visual range, the topological information set (such as connection, holes, nodes, branches and so on)of the image Y can be expressed as $Y = \{y_1, y_2, \dots, y_m\}$, where any element of Y can be obtained by the compression and expansion of the corresponding element of object X (in order to discuss conveniently, every set is supposed to have m elements, but that not means different sets have the same number of the elements). The topological information set of the object X is expressed as $X = \{x_1, x_2, \dots, x_m\}$.

Suppose the power set of Y (the collection containing all subsets of Y) is Ψ_Y ,

$$\begin{aligned} \Psi_Y = & \{ \{y_1\}, \{y_1, y_2\}, \dots, \{y_1, y_2, \dots, y_{m-1}\}, \\ & \{y_2\}, \{y_2, y_3\}, \dots, \{y_2, \dots, y_m\}, \\ & \{y_3\}, \{y_3, y_4\}, \dots, \{y_3, \dots, y_m\}, \\ & \dots, \{y_{m-1}, y_m\}, Y, \emptyset \} \end{aligned}$$

Suppose the power set of X is Ψ_X

$$\Psi_X = \{ \{x_1\}, \{x_1, x_2\}, \dots, \{x_1, \dots, x_{m-1}\}, \\ \{x_2\}, \{x_2, x_3\}, \dots, \{x_2, \dots, x_m\}, \\ \{x_3\}, \{x_3, x_4\}, \dots, \{x_3, \dots, x_m\}, \\ \dots, \{x_{m-1}, x_m\}, X, \emptyset \}$$

Proposition 3.1: Ψ_Y is the topology of the topological information set of the image Y , then (Y, Ψ_Y) constitutes a discrete topological space.

Proof: Because Ψ_Y , the power set of Y , contains all subsets of Y , obviously Ψ_Y satisfies three topological theorems as follows:

1. Both Y and \emptyset are in Ψ_Y ;
2. The union of random number of any subcollection of Ψ_Y is in Ψ_Y ;
3. The intersection of limited number of subcollection of Ψ_Y is in Ψ_Y .

Therefore, Ψ_Y is the topology of the topological information set of the image Y , and (Y, Ψ_Y) constitutes a discrete topological space, thus the proposition is proved and established.

Similarly, (X, Ψ_X) also constitutes a discrete topological space which denotes the topological space of the topological information set of the object X .

People can not only see the whole object, but also a part of it, thus the elements of Y can combine into any different subset which means the whole information of the image, the whole information of a partial image, the whole information of certain feature of the image (Although the integration of features in image is not separated, the combination among the elements of Y still can be characterized)

Proposition 3.2: When other conditions remain unchanged relatively, as the distance between the eyes and the object changes, the topological space of the topological information set of the image (Y, Ψ_Y) and (X, Ψ_X) which is the topological space of the topological information set of the object X have the same homeomorphism.

Proof: Y is the image of the object X generated on the retina, the topological space (X, Ψ_X) of the object X is the direct foundation of the topological space of the image Y . (Y, Ψ_Y) is the result of (X, Ψ_X) converted by human visual perception system. When Other conditions remains relatively unchanged and in the range of human visual perception, as the distance between the eyes and object changes, the image Y is just the compression or expansion of the object X , but the topology of it has not changed. As shown in Fig 6, for any element (that is topological properties, such as connectivity, the number of "hole" and so on) in the image Y there is a unique corresponding element in the object X . For example the elements y_1 and y_m of the image Y respectively corresponds to the elements x_1 and x_m of the object X .

Therefore, there exists a bijective correspondence between the object X and the image Y
 $f: X \rightarrow Y$

The mapping direction indicates the relationship between the reason and the result. Whereas, there also exists a bijective correspondence from the topological information set of the image Y to the topological information set of the object X

$$f^{-1}: Y \rightarrow X$$

Moreover, f and f^{-1} are continuous. According to the definition of the homeomorphism in topology, f is the mapping of homeomorphism. As a result, the topological space (Y, Ψ_Y) of the image Y and (X, Ψ_X) which is the topological space of the object X have the same homeomorphism. So the proposition is proved and established.

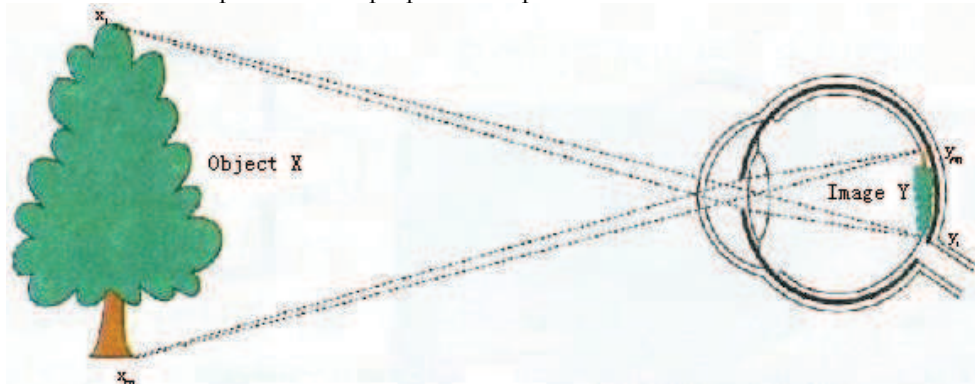


Fig. 6. The relationship between the object and its image

Proposition 3.3: The size constancy of visual perception has the property of topological invariance.

Proof: suppose the smallest distinguishable image (image of object X generated on retina from the farthest distance) is the topological space Y_1 (proved in Proposition 3.1 and corresponding to the of Figure 7(c)), as the distance between the eyes and the object changes, different topological spaces $Y_2, Y_3, \dots, Y_i, \dots$, are generated on the retina, Y_n is the topological space of the biggest image (as shown in Fig 7(b)). From proposition 3.2, when other conditions remain relatively unchanged, as the distance between the eyes and the object changes, the topological spaces of the images have the same homeomorphism as the topological space of the object X . So $\forall i \in \{1, \dots, n\}$, Y_i and X have the same homeomorphism, which means $Y_i \cong X$. Also because the property of homeomorphism has the transitivity, every Y_i is has the same homeomorphism with each other, which means for $\forall i, j \in \{1, \dots, n\}$ all satisfies $Y_i \cong Y_j$. Thus the topological spaces $Y_1, Y_2, \dots, Y_i, \dots, Y_n$ have the same homeomorphism, indicating that the size constancy of visual perception has the property of topological invariance.

In figure 7, the above are the images and below are the topological structures corresponding to its image. From the figure, (b) and (c) can be perceived as a cuboid because they have the same topology. For image (a), the size of the image is expanded seriously, only a few partial information about the object can be seen, therefore the object cannot be recognized for not

enough information available; in figure (d), the image is compressed so excessively that we are not sure whether it is a cuboid, cylinder or a small piece of others. There does not exist a bijective correspondence between the topological spaces of the (a), (d) and the topological spaces of the (b), (c), therefore, they do not have the same homeomorphism as the topological spaces of the (b) and (c).

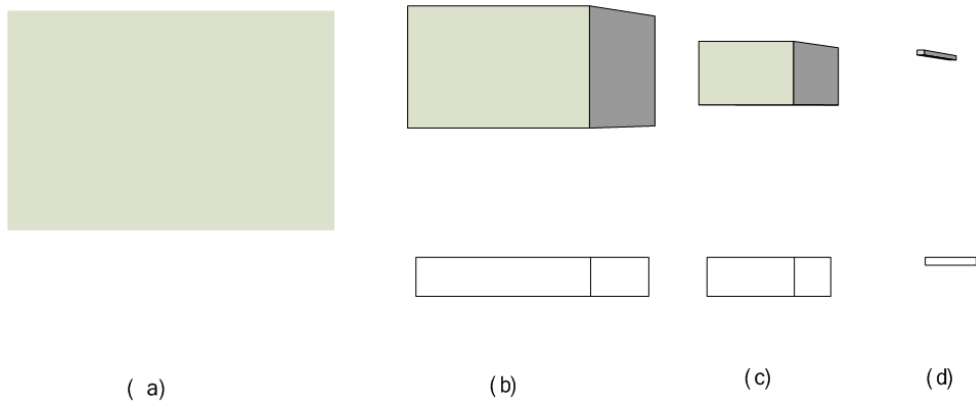


Fig. 7. The Size constancy

Thus, it is concluded that when the sizes of the images change, as long as the topology of the images has not been changed, people can perceive them as the same object, that is to say, the size constancy can perceive the objects which make the topological transformation as the same object.

Second, shape constancy :

As mentioned above, shape constancy is a kind of information procession which mainly as the angles between the eyes and the object change, the shapes of the images on the retina change, but in our minds we perceive different shapes of images as the same object.

The explanation of the shape constancy is shown in Figure 8.

Suppose Y_1 is the topological space of the distinguishable image (the image generated on the retina when the eyes are at the bottom of the object), different topological spaces $Y_2, Y_3, \dots, Y_i, \dots, Y_n$ of the images are generated as the angles between the eyes and object change, Y_n is the topological space of the image generated on the retina when the eyes are at the top of the object. Obviously the relationship between the topological spaces the images of Y_1 and Y_i ($i=2, \dots, n$) are squeeze and stretch. As the angles between the eyes and the object changes, part of the object are squeezed, while other part of the object are stretched, as shown in the figure 8. But as constancy, which means that, at the range of the human visual perception, each side and each part of the object keep the original characteristics well, without adhesion or tearing. Therefore the topological spaces $Y_1, Y_2, \dots, Y_i, \dots, Y_n$ have the same homeomorphism, namely the shape constancy of the human visual perception has the property of the topological invariance, the method of the proof is just the same as that of the size constancy.

In Fig 8, the above row are the images and below are the topological structures corresponding to its image. From the figure, (a), (b) and (c) can be perceived as a cuboid, because they have the same topology. (b) and (f) have the same topology different with

others, we cannot perceived them as a cuboid, maybe we will perceived them as folded piece of a rectangle. We may perceive (g) as a piece papper.

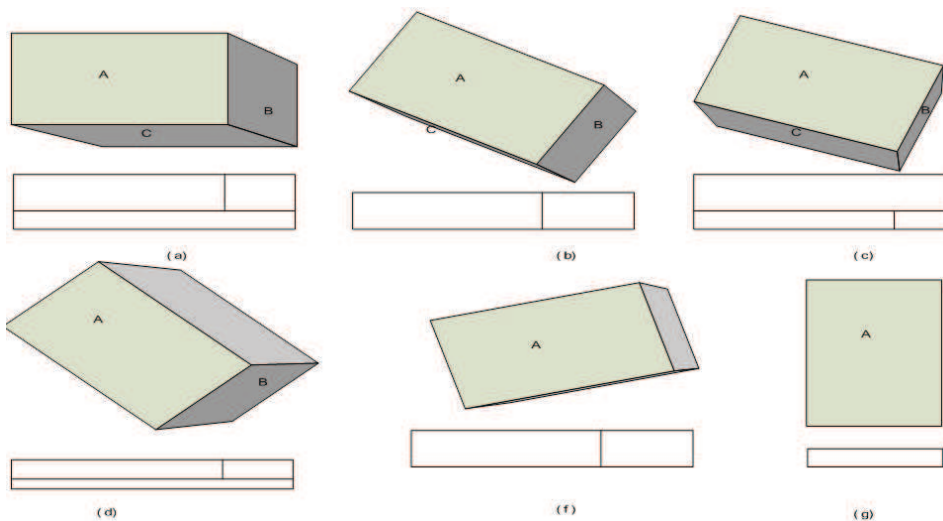


Fig. 8. The shape constancy

Therefore, it can be concluded that when the images' shapes of the object change, as long as the topology of the images has not been changed, we can perceive them as the same object, in other words, the shape constancy can perceive the objects making topological transformation as the same object.

From the above analysis, we can see that the constancy in size and shape of the visual perception can be described by topological invariance. As long as the topology of the images keeps unchanged, we can perceive the images that make topological transformation as the same object at the range of the human visual perception.

In conclusion, we can believe that the constancy of the human visual perception can be described by the topology invariance, that is, people can perceive the objects which are changed by topological transformation as the same object.

3.3 Composition of patterns

From the theory of the component recognition, the principle which the components constitute objects in the real world can be understood like this: there are countless type of materials in the world, but the types of the chemical elements composing the materials are just a little more than 100. Various colors in the world can be composed by the three colors of red, green and blue. All buildings, no matter how grand and splendid they are, are composed by the foundation, beams, column doors, walls, floors and the ceilings. A variety of delicious food can be cooked by just a few spices of the oil, salt, sugar, vinegar, soy sauce and so on.

Music is wonderful, and countless musical works are composed by the seven notes of 1,2,3,4,5,6,7 and their changeable sound. The number of the English words is very huge, moreover in the rapid development, but they are just composed by the 52 letters including

26 uppercase and 26 lowercase letters. As we know the number of figures in various aspects of the world is endless, but in the decimal system these figures are constituted by 0 to 9 altogether ten figures; in the hexadecimal 0 to 9 ten figures and A to F six letters are needed; in the binary only two symbols 0 and 1 are needed.

Therefore we can take it that Biederman's theory of the components recognition reveals the pattern of the real world's construction: all the objects of the world are composed by a few components, that is to say, all the objects can be decomposed into certain components.

Biederman's theory holds that the limited components have almost infinite combination, thus compose almost unlimited objects. This conforms to Chinese philosophy of "one lives two, two lives three, and three lives everything". In terms of the geometry, the objects can be fully described by the geometries, because various spatial relationships among the geometries have infinite combination. The same components can form different objects through different combination. The English words "on" and "no", "dear" and "dare", "hear" and "hare" respectively have the same letters, but when combination is different, the words composed are completely different.

We call such a combination of the objects structures, which is a very important concept in pattern recognition. The definition of the structure is the organization of each part of the object[19].

In pattern recognition, the structure is the combinational relations between the object and its' components, and the combination of their components. For example, the English word "structure" is composed by s, t, r, u, c, e six letters, the order of the arrangement is s-t-r-u-c-t-u-r-e. In organic chemistry, Methane (CH_4) is composed by the two elements of Carbon(C) and Hydrogen(H), and the structure of the CH_4 is a regular tetrahedron, that the carbon atom is in the center of the tetrahedron, while four hydrogen atoms are in the vertices of the tetrahedron, as shown in Fig 9.

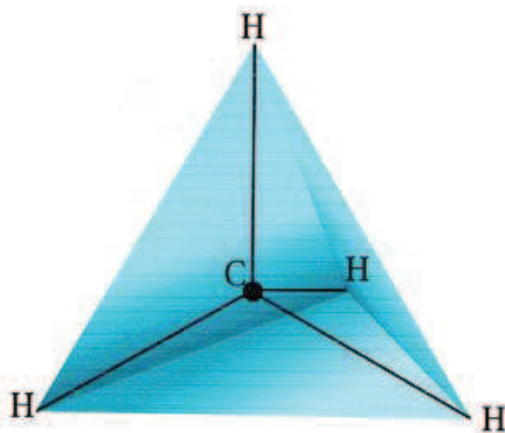


Fig. 9. The structure of the methane

The structure of objects, has its own laws, which cannot be constructed arbitrarily. In English, the letters have certain regulations of arrangement to construct the words, for example we cannot arrange w, v, u and some other letters together repeatedly. In appearance of the human face, there are not more than two eyes, only one nose and only one

mouth, moreover the distribution of them is that the eyes are at the top of the face, the nose in the middle, and the mouth at the bottom.

Now we use the theory of set to describe the construction of the object:

Definition 3.1 The objects which have the same characteristics in some aspects compose a set, denoted as O .

Definition 3.2 The elements in the object set O are called object or pattern, denoted as m , that is $m \in O$.

Definition 3.3 For any $m \in O$, we can decompose it to one or several components according to a certain regulations, then the components form a set which is recorded as C_i (i is the natural number)

Definition 3.4 For any $m \in O$, we call the mutual relationship of the components structure which can be denoted as S or $S(C_1, C_2, \dots, C_n)$.

For example, with a certain structure the word "study" is composed by the letters "s, t, u, d, y", which can be denoted as follow:

$$\text{Study} = S(s, t, u, d, y).$$

All patterns in the object set can be decomposed into a component or more according to a certain structure, that is, all the objects in the set can be composed by the several components according to a certain structure.

3.4 The relationship between the prototype and the components of objects

The theory of prototype believes that the storage in the long-term memory is the prototypes, rather than the templates corresponding to the external patterns. What is the prototype after all? Prototype is the internal characteristic of a class of objects and also the general characteristic of all individuals of a class of objects, but not the internal copy of a specific model. It is considered by component recognition that objects are composed by a number of basic shapes or components, that is to say, are composed by the geometries.

From the above two theories, it appears that there is a difference in the content of their researches: the theory of the prototype matching is to study the human brain for perceiving the outside world, while the theory of the component recognition is to study the composition of the objects. However, the two theories are to study human's pattern recognition, is there any relationship between them?

Prototype, the general characteristic of all individuals in a category or area, reflects a class of the objects' basic characteristics, is the generalization and abstraction of the object's characteristics. As analyzed in the preceding chapter, an object is constituted by some elements that are concrete and determined under its structure. Therefore, there exists a process from general to determine and from abstract to concrete between the prototypes and the components.

We might understand prototype like this: for an object set of English word, and for any word in the object set, all of them are constituted by one or more of 52 symbols which has 26 uppercase and 26 lowercase letters. With 26 uppercase and 26 lowercase letters, the 52 symbols are the generalization and abstraction of the components of all the English words. No matter how many English words and various fonts there are, they can be constituted by some letters changing in size and shape. Therefore, English words' uppercase and lowercase letters are the prototypes of this object of English words. The sizes and shapes of prototypes

are fixed, the matches between them and the components with various size and shape can be realized by topological mapping. The deduction of the concept is shown in Figure 10. In Fig 10, the prototype set P has 52 elements including 26 uppercase and 26 lowercase letters, while all the English words compose the set of the objects. The elements in the set of the objects are specific English words "Pattern Recognition PATTERN RECOGNITION", which are composed by elements of the prototypes "R, P, A, T, E, N, C, O, I, G, e, c, o, g, n, I, t". The elements' size and font in the prototype set are fixed, the prototypes can match each kind size and font of the components through the topological transformation.

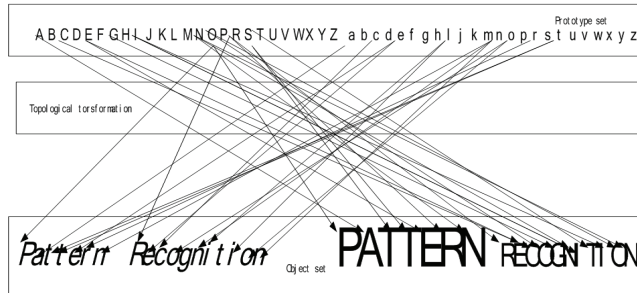


Fig. 10. The relationship between the prototypes and the components of objects

It can be described from the perspective of topology:

Definition 3.5 A set without any repeated elements abstracted from the component collection C of all the objects M in a kind of object O, is called the prototype of this kind of object O, denoted as P,

From the definition 3.3 and 3.5, all objects M in every class of object O can be decomposed, under a certain rule, into some components C, from which a set abstracted without any repeated elements becomes prototype P, as shown in Fig 11.

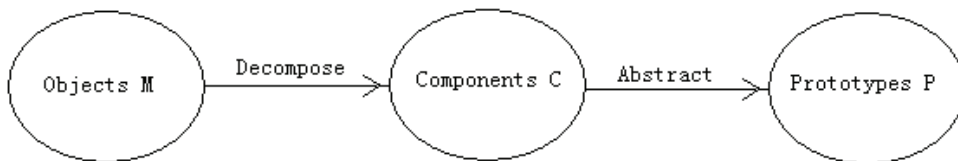


Fig. 11. Objects-Components-Prototypes

In the example above the specific objects are the following English words "Pattern, Recognition, PATTERN, RECOGNITION", if set $M = \{Pattern, Recognition, PATTERN, RECOGNITION\}$, then M can be decomposed into the component collection $C = \{P, a, t, t, e, r, n\}, \{R, e, c, o, g, n, i, t, i, o, n\}, \{P, A, T, T, E, R, N\}, \{R, E, C, O, G, N, I, T, I, O, N\}$. The differences of the components among the R, R and R are only size and shape, there is no difference on topology among these three elements, so these three elements can be abstracted as a prototype R. Therefore, the prototype set P can be abstracted from the component collection C, that is the prototype set $P = (R, P, A, T, E, N, C, O, I, G, e, c, o, g, n, I, t)$.

Thank You for previewing this eBook

You can read the full version of this eBook in different formats:

- HTML (Free /Available to everyone)
- PDF / TXT (Available to V.I.P. members. Free Standard members can access up to 5 PDF/TXT eBooks per month each month)
- Epub & Mobipocket (Exclusive to V.I.P. members)

To download this full book, simply select the format you desire below

