

THE MODEL OF HUMAN VISUAL LONG-TERM MEMORY WITH ABILITY FOR GENERALIZATION.

R.M. Granovskaya

Leningrad University, USSR

Summary

A model of human visual long-term memory with ability for generalization has been synthesized. Perceptive structure of the system may be either parallel (multilayered retina) or successive (a network scanning a contour by tracing).

The synthesized model of retina performs a number of transformation of information as in a visual analyzer - such as intensification of contrast between an object and its background, contour or outline extraction, reaction to the movement of an object (in a particular direction at a certain velocity).

It is the code tree that is used as a structure for storing codes in long-term memory. This structure explains the dependence of the response time upon the number of alternatives and allows studying the dynamics of the following trajectory (path) reduction occurring during learning process.

A model of human long-time memory presented as a structure of a tree consisting of neuron models is described and such processes of information transformation in it as recording, storage, ge-

neralization, reading, disposition set to the given type of signals are considered. The perceptive structure of this system may be either successive (scanning) or parallel (multilayered retina).

There are various kinds of scanning:

a) successive, ranked, one-channel tracing of the contour and the formation of primary codes. b) distant, uneven tracing on characteristic points of objects and the formation of secondary codes. c) transition to the parallel multichannel scanning through the retina.

During the tracing the contour is divided into segments according to a definite rule. The number of segments is determined by the contour configuration only. Each segment is characterized by a ranked group of the primary signs ($\alpha, \beta, \gamma, \delta, \lambda, \varepsilon, \eta$) which are functions of its curvature. The primary code is a table in which the number of lines and columns corresponds to the number of signs used and the number of segments of the contour.

The use of the one-channel contact tracing brings us to the necessity of introducing information about the entire perimeter in the form of a definite simple sequence, which reduces the rate and increases the number of limitations on the condition of perception. Digression from the contact method of percep-

tion made it possible to elaborate the corresponding extended secondary signs of the object and to form the secondary code, out of them. The secondary codes are more generalized and shorter than the primary ones (they are one line instead of a table as is the case with the primary codes) and they simply correspond to an object, the sequence of presenting the signs being of no importance. The use of the secondary codes in the process of identification made it possible to trace characteristic points of an object irrespective of the sequence of the tracing, which reduces the number of limitations on the object under perception and makes it possible to place it into the field of vision irrespective of sequence. Besides this the characteristic parts of an object may be presented, not an entire object.

The secondary codes are invariant with respect to orthogonal, similarity, affine transformations of the contour and transformations of the central projection within definite bounds as well.

The code tree is used as the long-time memory structure for the primary and the secondary codes. This structure explains to us how the time of the reaction depends on the number of alternatives, makes it possible to fix how often separate events occur, and to examine

the dynamics of the tracing trajectory reducing while training (learning).

Multichannel parallel introduction of information about visual objects into memory is performed through retina in the model. This means that a certain retina structure will provide for the investigation of an object in a definite sequence, division of the information thus obtained into pieces and its investigation by these pieces.

Assuming that sensitivity of perceiving elements of the visual system changes so that at every moment sensitivity of only one group of elements exceeds the threshold while that in other groups does not reach it (the threshold) only those parts of an object are picked out simultaneously that are projected on the elements of high sensitivity. Hence, an object is picked out successively by certain parts, and every part is picked out parallel to working elements. In this case we have the system working parallel-successively.

In case we have parallel, not successive introduction of information on a uniform retina it may be demonstrated that while the thresholds of the elements of retina change according to a periodic law, the information about the object is transferred from the retina to memory successively by parts, the nature

of succession being dependent on the object properties. However dependence is different from that when the contour is traced. It means that the stabilized (motionless with respect to the retina) object can be perceived and introduced in memory, in these conditions, too.

Transition to such a type of perception is also a kind of tracing the closed cycle which is set by a period of threshold changes, the dynamic process of threshold changes being independent on the movements and may be set centrally, which quickens the process of perception. Periodic changes of the threshold make it possible to obtain a series of codes of one and the same object (ranked according to the law of element sensitivity changes). Formation of the group of codes of one object, the group connected with the inner period of the perceiving elements sensitivity changes without the outer transformation of the object enables one to make an approach to the explanation of certain aspects of the constancy of perception mechanism.

Thus, when the visual system works in a successive-parallel way the movement of the tracing system of the eye may be replaced by the dynamic regime of the perceiving neurons' sensitivity changes.

The model of the retina is a structure consisting of homogeneous elements. Each

element is a multi-layered network. The first layer is the receptive field of an element. The second one consists of novelty neurons. This layer picks out brightness changes in the receptive field and, thereby, the object's border. The third layer consists of summing neurons. The line curvature is detected irrespective of the line orientation in this layer, i.e., the element in this layer yields the output proportionate to the line curvature in the receptive field. The fourth layer consists of difference neurons and responds to the movement of the picture and performs selection of the time-space code, picking out certain gradations of the curvature (Fig.1,2).

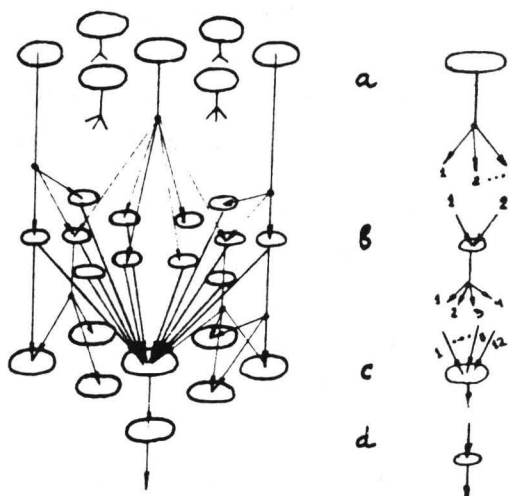


Fig.1 Model of retina main element.
(a) The first layer - the receptive fi-

eld of an element. (b) The second layer - novelty neurons. (c) The third layer - summing neurons. (d) The fourth layer - a difference neuron. The types (b)-(d) neurons equations are shown in the figure 6 of the report "The model of Human Short-Term Memory" by I.Ja.Bereznaya and R.M.Granovskaya.

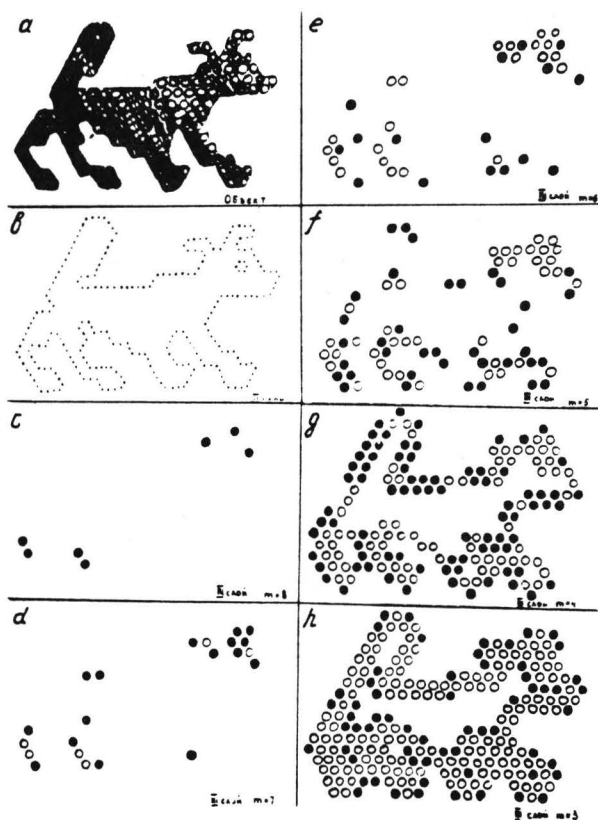


Fig.2 Object perception by the model of retina.

(a) The first layer - an object. (b) The second layer. (c)-(h) The third layer at successive time moments t_3 - t_8 , when neurons threshold value changes from $m=8$ to $m=3$.

Thus codes of object pictures at retina output are invariant with respect to all the orthogonal transformations, namely, a turn, parallel transition of a picture and mirror reflection.

The fact that there is no division of the neuron receptive field into exciting and inhibiting zones is the specific feature of the given retina model.

The offered retina model makes it possible to illustrate the psychological phenomenon of illusionary movement which consists in the perception of the seeming movement of one object, although two objects are presented successively at different points. The reaction to the signal change either in space or in time is the main principle of how all the layers of retina operate. The operation of the given retina is considered as a model of separate stages of visual objects perception (Fig.3).

The memory model is either three- or two-dimensional code-trees, one- or multi-layered, dichotomous or K-tomous, built of homogeneous or heterogeneous joints and elements: what is common about them is a structure resembling a tree. The usage of the code tree as the main structure of a long-time memory has great advantages over other types of memory arrangement, as for example the teaching matrix. Such structure ma-

kes it possible to use fewer elements than when codes are stored as an infinite sequence, which increases the rate of arriving at a decision when the code is searched in memory.

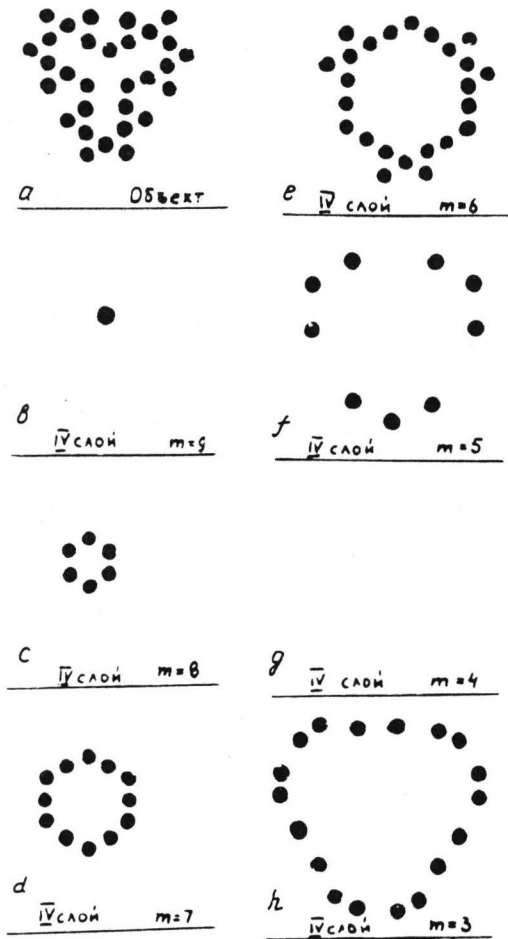


Fig.3 The illustration of the psychological phenomenon of the illusory object movement.

(a) The retina first layer - an object.

(b)-(h) The retina fourth layer at successive time moments when a neuron threshold value descends from $m=9$ to $m=3$.

The process of recording in the offered memory model presented as code trees is arranged so that a successively ranked set of object signs arrives at the input either directly from the outer input or from the output of the short-time memory. Each of the element after missing a certain code is inhibited while the code continues to be recorded in the upper layers of the tree. This results in the successive recording and reading of the codes.

In a situation like this a spatial code tree that makes it possible to reduce the number of alternatives in the tree joints is used for the storage of primary codes.

A two-dimensional code tree is also used as a structure for the storage of secondary codes. Each gradation in such a tree will have its corresponding branch, each code - its separate way, each sign - its layer. The identical codes will have their identical way, the repeating codes do not alter the tree configuration.

Accordingly, the tree will have as many layers as the number of signs used and as many ways as the number of different codes arriving at the input of the memory system. Each way of the tree is ascribed its weight proportionate to the number of the corresponding codes

that arrived at the input of the memory system.

The only way corresponds to the full code arriving at the tree input. The end of the way is connected with the corresponding index. Besides this index, indexes of all the joints lying in the given way can be counted. In case we have several ranks of the code arriving at the input, note the full code reading may be performed with but a certain degree of relativity (Fig.4,5)

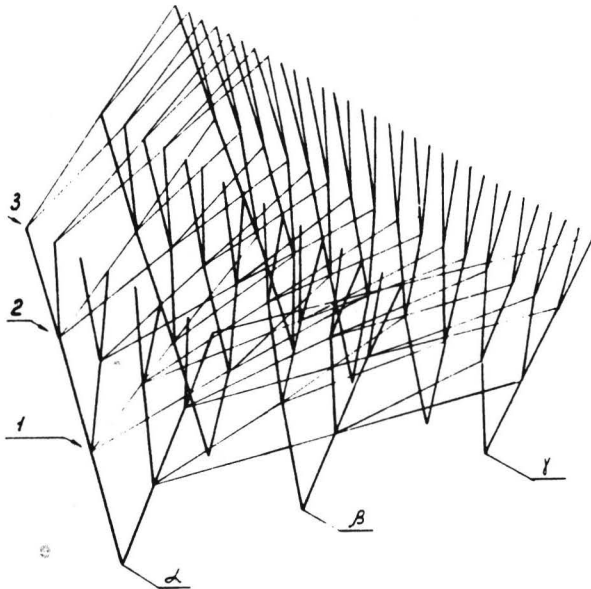


Fig.4 A space code tree formed when successive one-channel tracing of the contour (primary signs α, β, γ): α, β, γ - vertical plane code trees, 1, 2, 3 - horizontal plane code trees.

The tree structure provides for the only descent (simple identification) from any joint down to the tree foot which

le the reading takes place. If signs are different as to the number of gradations, the number of elements of the tree will depend on the succession in which signs are to be applied. The number of elements will be the smallest if the signs are applied so that the number of gradations increases. Omission of non-ramifying ends of ways in the full tree can serve as the following step to economize elements.

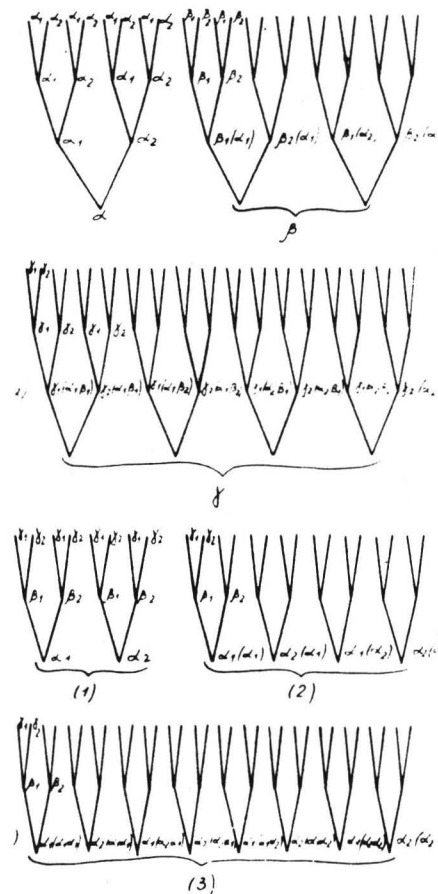


Fig.5 Aplane code trees ($\alpha, \beta, \gamma, 1, 2, 3$) that form the fig.4 space code

tree.

When the arrival of signs in the code is ranked, the memory structure is economized. First, the arrival of the code of the same object with a larger number of signs does not require additional code recording; second, the ranking of signs according to the number of gradations or to their importance makes it possible to reduce the number of steps along the tree so that the identification is with the same degree of accuracy, i.e. the ranking is advantageous as to the process rate.

When the ranked system of signs is used, the tree structure enables one to increase their number while the codes already in storage are not changed. This enables one to analyse objects and record codes with any degree of accuracy needed, moving from the tree foot up to its summits as it is relevant in some cases to adopt decisions at the expense of a rough classification and at the same time to reduce the time of decision without going up the tree. Such structure makes it possible to perform a simple recording and code reading by its index which is compared with one of its summits. Recording and reading are performed without address in this structure, associatively (in so far as the code has no special address group) and the code

arrangement on the tree is determined by the code properties only, the structure of its signs.

While reading, a complete shuffling of the codes in storage is not performed: the needed code or a group of nearby codes are read immediately, which quickens the process of reading. The time of selection does not depend on the bulk of the information recorded, being determined by the number of layers in the tree only.

The measure of depth of the code in storage has been introduced. It is presented as weights of those branches and joints of the tree which store information of the given code. The weight is connected with the changes of synaptic coefficients of those elements which the codes traverse in the process of teaching.

Weights accumulation in the tree structure is compared with the frequency of occurrence of the given code at the system input. If sequences with the maximum weight are read out off memory they will be the codes that have been most often presented in the history of the system.

It has been shown in the model that the optimum uneven strategy of the potential hypothesis check-up can be ela-

borated on the basis of the statistics in the tree accumulated in the process of teaching and the trajectory of the object investigation changes as the result of this. Thus, the brain produces a sort of reflection of an object which is expected to be on the retina; and this brings about changes in perception and speeds up the process of identification.

Arrival of information about new objects brings about changes in weight and in pseudoweight as well, i.e., there is accumulation of information in the memory structure not only about the codes that arrived earlier (weight) but also about these ones that are similar to them according to a certain criterion of similarity, i.e., codes are generalized by their similarity, the similarity being determined by the nearness of the signs arrangement in the memory structure or well by the nearness of signs in the ranked sequence when they arrive at the memory input (pseudoweight).

The usage of pseudoweight makes it possible to reduce the period of teaching (learning) through the reduction of the teaching sequence and to teach the system of memory with the help of the most characteristic representatives of the class, only bearing in mind that a sort of intermediate relief stretches automatically between the excited parts

of memory. Such a system starts responding to objects that were not presented earlier but are part of a class of such objects and are similar to them according to certain **criteria**. The system responds to the objects as if they were known to it.

The usage of pseudoweights makes it possible to optimize (reduce) the teaching sequence according to the criterion needed and thereby reduce the process of teaching the system the given dictionary of objects.

The introduction of weights and pseudoweights makes it possible to approach the modelling of such important psychological process as the formation of a unit determined by the immediate prehistory.

The usage of pseudoweights also makes it possible to extrapolate properties of nearby objects i.e. to solve the task before its exposition is over or even by way of an imaginary experiment with it.

Assuming that there is a certain tree structure, homogeneous as concerns its properties before the teaching, the structure will grow more and more heterogeneous in the course of teaching in so far as the changes in the distribution of weights of branches and joints go. The greater the number of different codes that arrived at the given memory

structure, the greater the number of tree joints used, i.e., the history of the system gradually goes up the initially homogeneous tree, i.e., it sort of grows. The heterogeneity thus appearing is isomorphous to the group of the passed codes both by the nature and the area it spreads over and makes it possible to restore the history of the system with a certain degree of accuracy later on.

After the memory structure had passed through a certain sequence of codes it is a heterogeneous system consisting of ensembles of functional elements with the rate of processes varying in different parts of the system and the signals pass from one joint to another with the varying probability. Values of the processes rate in the ensembles and the distribution of weights of the branches that determine the branch the process will spread through are determined by the prehistory of the system in the process of teaching only. The usage of a special system independent of the input sequences to control the thresholds of the functional elements of the tree in the model makes it possible to approach the explanation of how the importance of the task and its emotional colouring influence the narrowing of the solution search zone.

When such system is launched the way

the given code in the tree will follow will be determined not only by the prehistory of the memory system but also by the changed sensitivity of the upper joints of the tree. In this case the code will follow another branch and will bring about another solution (it is identified with another index).

The measure of novelty of the solutions thus obtained is determined by the number of potential prolongations from the given joint. If there is a group of ramifications the weights of the sum of prolongations being equal, the changes in the sensitivity of one of the upper layers will essentially narrow the search zone, i.e., the task in this particular case comes out as a factor of the search directness and determines the formation of the unit connected with the signal importance.

Reduction of the identification of printed letters of the Russian alphabet has been considered as an example of reduction of the process of identification. Primary codes concerning signs $\alpha, \beta, \gamma, \delta, \lambda$ are patterns of letters by the successive scanning. It has been shown that the greater the number of signs used the smaller the number of groups of the primary code and the corresponding part of the contour required for the simple identification. When only one

sign α is used the alphabet falls into 18 groups, when signs α and β are used there are as many as 28 groups. All the letters are mutually separate with a larger number of signs. The average number of intervals per letter necessary for the identification will be 13,3, in case all the five signs are used the number of intervals will be 4 (Fig.6).

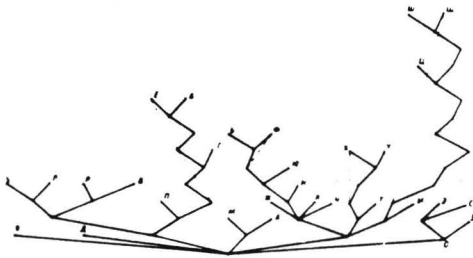


Fig.6 Russian letters code tree when letters contours are traced successively.

Further reduction of the process can be achieved at the expense of transition from the continuous tracing along the contour line to the uneven one.

Direction and value of the leap are determined by the hypothesis put forward on the basis of the analysis of the upper part of the tree structure. The leap must result in the hypothesis check-up (Fig.7).

When information about letters of the Russian alphabet is introduced in memory with the help of retina all the letters are subdivided into groups whose number depends on the number of gradations of

the retina elements sensitivity. It has been shown that if the number of gradations is six. All the letters are discerned by the retina. The smaller the number of gradations the rougher the classification i.e., several letters find themselves in one and the same group (Fig.8).



Fig.7 The reduction of Russian letters contour tracing during identification concerning signs α , β , γ , δ , ϵ .

The organization of memory according to the principle of the code tree makes it possible to perform the object identification through successive approaches while going from the foot of the tree to one of its summits. The change of hypothesis from a more general to more particular one up to the precise identification takes place in the process of such a movement.

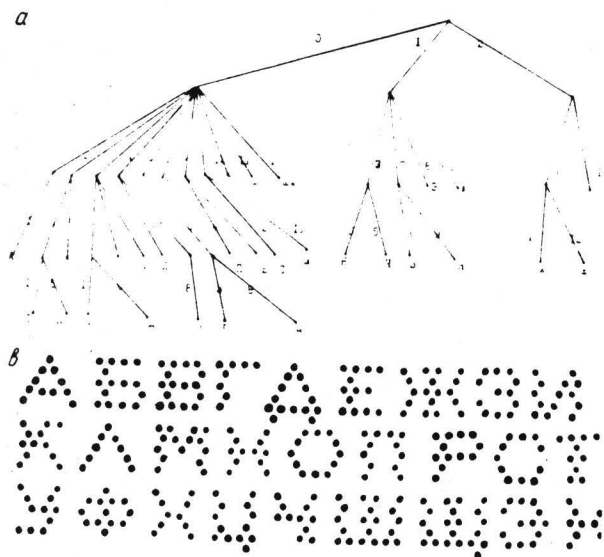


Fig.8 (a) Russian letters code tree after parallel multichannel scanning through the retina model when the model fourth layer neuron has 6 sensitivity change ranks. (b) Russian letters on the retina model.