DOCUMENT RESUME

ED 415 840 IR 018 691

AUTHOR	Metallinos, Nikos
TITLE	The Transformation of the Biological Precepts into Mental Concepts in Recognizing Visual Images.
PUB DATE	1998-00-00
NOTE	28p.; Paper presented at the 1998 Paper Competition of the Division of Production, Aesthetics, and Criticism of the Broadcast Education Association (Las Vegas, NV, April 3-6, 1998).
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Aesthetics; Biological Sciences; *Cognitive Processes; *Communications; Films; Foreign Countries; *Mass Media; Perception; Photographs; Programming (Broadcast); *Television; Television Research; Theater Arts; *Visual Arts
IDENTIFIERS	Canada

ABSTRACT

Total comprehension and understanding of visual images (particularly moving images) are the result of a series of complex biological (brain) and mental (mind) processes and activities. Whereas perceptual psychology and neurophysiology are among the two main academic disciplines that explain the functions performed by the organs of visual and auditory perception (eyes, ears, brain), cognitive and behavioral psychologies are the main academic disciplines that explain the mental activities, processes, and functions of the mind. In this paper the transformation of the biological precepts into mental concepts is discussed as they relate to recognizing and understanding moving visual images. Specifically, this paper reviews the various biological and mental functions of the human brain as they relate to moving images, discusses how visual precepts (codified visual bits) are transformed into visual concepts (holistic visual units), and provides suggestions as to the construction of moving images, particularly televised images. (Contains 31 references.) (Author)

The Transformation of the Biological Precepts into Mental Concepts in Recognizing Visual Images

by

Nikos Metallinos, Ph.D.

Paper presented at the 1998 Paper Competition of the Division of Production, Aesthetics, and Criticism of the Broadcast Education Association

April 3-6, 1998, Las Vegas, Nevada

Concordia University

Department of Communication Studies

7141 Sherbrooke Street West

Montreal, Quebec H4B 1R6

Canada

Abstract

Total comprehension and understanding of visual images (particularly moving images) are the result of a series of complex biological (brain) and mental (mind) processes and activities. Whereas perceptual psychology and neurophysiology are among the two main academic disciplines that explain the functions performed by the organs of visual and auditory perception (eyes, ears, brain), cognitive and behavioral psychologies are the main academic disciplines that explain the mental activities, processes, and functions of the mind. In this paper the transformation of the biological precepts into mental concepts is discussed as they relate to recognizing and understanding moving visual images. Specifically, this paper (a) reviews the various biological and mental functions of the human brain as they relate to moving images, (b) discusses how visual precepts (codified visual bits) are transformed into visual concepts (holistic visual units), and (c) provides suggestions as to the construction of moving images, particularly televised images.

Introduction

The merging of science and the arts is today more demanding than ever before. What was previously a mere suggestion by and desire of scientists and artists (Behnke, 1970; Metallinos, 1983), is now a necessary practice. Computer science has bridged the gap between these traditionally disassociated disciplines to the extent that scientists are becoming artists and artists scientists. This is evident in the area of visual communication media arts, primarily the electronic media arts.

The artistic approach in the creation of electronic communication media artifacts such as film, television, and computerized images, which was based--for the most part--on the humanities and social science--relying heavily on the content of the message--today is interrelated with the scientific approach, mostly concerned with the technology of the media. The various instruments, special materials, and established techniques used to construct the electronic media messages are more important than the verbal content of the message (Tarroni, 1979). The cameras, lights, and microphones, along with the framing, lighting, sound, and editing equipment and techniques, are media technologies that require considerable scientific and artistic knowledge and experience to be better handled and generally more effective. The computerization of all media areas (i.e., instruments, materials, and techniques) and all production levels (pre-production, production, and post-production) is a vivid example of the interrelation of scientific and artistic requirements. The physical or technological properties of the electronic media require basic scientific knowledge and proper synthesis of all production elements requires artistic inspiration and emotional literacy. When electronic communication media programs in general, and television production in particular, are not the result of this scientific and artistic duality, they are usually mediocre, uninspired, and, for the most part, mundane. Scientific knowledge and artistic sensibility, that is, intellectual and emotional literacy, are not only desirable, but main prerequisites for the construction and creation of electronic media products, particularly visual communication media productions. Consequently, the need to study, understand, and apply the scientific findings by perceptual psychologists and neurophysiologists regarding the various functions performed by the organs of perception in general, and the brain and the mind in particular, is now greater than ever before.

Regrettably, visual communication media studies that acknowledge the scientific basis of visual, auditory, and motion perception and cognition are minimal or non existent (Metallinos, 1996), whereas in the area of composition of moving images the opposite is true. Traditionally, communication scholars mostly derive from and rely heavily on social science and humanities, which explain the plethora of media studies examining the content of their messages as the cause for their audience impact. The study of any medium, however, as an art form has to be holistic; that is, the researcher/investigator must consider equally, the message, the medium or form, and the target audience in their inquiries. However, this was seldom the practice in the past.

An area in the study of the visual communication media that requires scientific knowledge along with artistic sensitivity is the production, or construction, of visual images for maximum aesthetic impact and artistic effect. What should the producers/directors of visual communication media know regarding the processes involved in the perception and codification of visual images by the brain, and their subsequent decodification and recognition by the mind? If one knows how the bits of information contained in such visual images as television pictures (which incorporate sights, sounds, and motion) are transformed into recognizable images, one selects those particular visual, auditory, and motion elements that are readily recognizable and generally more interesting. The purpose of this paper is to examine how the various biological precepts--contained in visual images--are transformed into recognizable concepts, given that total comprehension and understanding of such images are the result of complex biological (brain) and mental (mind) processes and activities. Specifically, this paper (a) reviews the various biological and mental functions of the brain as they relate to moving images, (b) discusses how visual precepts (codified visual units) are transformed into visual concepts (holistic visual images), and (c) provides specific guidelines regarding the construction and artistic synthesis of moving images, particularly television pictures.

Brain and Mind Functions in Moving Image Recognition

Scientific studies, mostly by neurologists, optometrists, neurophysiologists, and psychologists, regarding the unique functions of each hemisphere of the brain, and the mental processes involved in the final recognition of objects, subjects, and events are in abundance (Bloom, Lazerson, & Hofstadter, 1995; Springer & Deutch, 1985). They have drastically increased during the last twenty years, mostly due to the advanced research instruments and measuring techniques heavily driven by computers (Haber, 1968; Martin & Venables, 1980). However, the transformation of visual, auditory, and motion inputs into recognizable entities remains a mystery, although the processes involved have been closely observed, theorized, and identified, mostly in the areas of patterns and shapes of objects and figures. As there is some confusion in the ways by which communication scholars and artists are using the terms stimulation, sensation, perception, and cognition, it is necessary to review them before examining the scientific theories on visual recognition and visual imagery.

Stimulation is the process by which the external phenomena, by eliciting their stimuli in various forms (i.e., electromagnetic waves), trigger the organs of perception (or receptors) and cause their response. The eyes, ears, nose, etc., first are stimulated. The ultimate purpose of stimulation is to trigger the organs of reception and cause their subsequent sensation or response. The response is determined by the degree of stimulation, and vice versa. Consequently, we divide the stimuli into effective or responsive, subliminal or unconscious, and ineffective or unresponsive (Murch, 1973). Constructors of visual images have the ability, through their visual messages, to stimulate, in various degrees, their viewers. Because the stimulus is the cause and the response is the effect, we must examine next, the term sensation.

Sensation is another name for the response, which is caused by the various degrees of stimulation. To sense is to respond, and to respond is to stimulate. The degree to which we are stimulated by the various phenomena such as visual images not only depends on the various levels of stimulation above, but is also determined by the physiological conditions of our organs of perception. Such organs are called channels of stimulation and they are: vision (sight), audition (hearing), gustation (taste), olfaction (smell), and tactile kinesis (touch and body position). Weak or defected organs of reception obviously do not respond, or they do not sense as effectively as healthy ones. Equally, untrained and inexperienced to various sensation organs do not respond the same as organs accustomed to the various sensations. These commonly known common sense factors are usually ignored by media scholars and researchers who measure audience responses to moving images, assuming, inaccurately, that all viewers or listeners sense, response, and perceive them uniformly, equally.

Perception is the process by which the channels of stimulation or organs of reception receive the physical sensations, become aware of them, and categorize and codify them. Perception, therefore, is the codification process of the stimulus/response (stimulation/sensation) of the environmental conditions that trigger the perceptual centers; it is the awareness of the objects and events of the environment through the physical sensations; it is the physical sensation codified, or semi-interpreted. Consequently, we can state that: to stimulate is to sense, to sense is to respond, to respond is to be aware, or to be able to distinguish and codify.

In the process of visual, auditory, and motion perception of images, perception goes as far as arranging the received stimuli into cohesive units, bits of information--classified and intensififed--and prepares them to be processed into the brain's special regions for decodification, interpretation, and inevitable recognition.

Cognition is the process by which raw data, the classified and intensified bits of information, are turned into holistic recognizable units. Cognition is synonymous to comprehension, recognition, understanding, interpreting. It means to be cognizant or knowledgeable of an object or event. Because to perceive is to be stimulated, to receive and to be able to codify raw data of information, to recognize is to be able to decodify, demystify, interpret the information process into holistic entities of completed information. Therefore, all perceptual processes are neurophysiological and biological activities of the eyes, ears, and brain, whereas all cognitive processes are mental activities performed by the brain and the mind combined. In perception we see, hear, taste fragmented bits of information. In cognition we see, hear taste, cohesive, unified information.

This distinction between perception and cognition, and the clarification of processes involved in each case is of paramount importance to the constructor of visual images--primarily moving images--as it will be evident in the forthcoming discussion of theories of visual recognition.

Although literary sources on the theories of visual communication are in abundance in the field of perceptual psychology, there are only a few in the field of communication. The traditional theories of visual recognition by Pinker (1988) and the sensual theories of visual communication by Lester (1995), representing the fields of perceptual psychology and communication respectively, will be examined herein.

Pinker (1988) was the first psychologist and cognitive scientist to point out that visual cognition is a dual process, physical representation and mental reasoning of the phenomena. This important distinction is explained by Pinker (1988) as follows:

> Visual cognition can be conveniently divided into two subtopics. The first is the representation of information concerning the visual world currently before a person...Visual recognition is the process that allows us to determine on the basis of

retina input that particular shape, configuration of shapes, objects, scenes, and their properties are before us. The second subtopic is the process of remembering or reasoning about shapes or objects that are not currently before us but must be retrieved from memory or constructed from a description. This is usually associated with the topic of visual imagery. (pp. 2-3)

The traditional theories of shape recognition, according to Pinker (1988), are: the template matching, the feature models, the Fourier models, the structural descriptions, the Marr-Nishihara, and the massive parallel models.

The Template Matching Theory states, in effect, that the retinal stimulation projected by the shape of an object matches, or is simultaneously superimposed by all the templates existing in memory, until the one which is closest to the retina will prevail, indicating the actual pattern present.

This is the simplest theory of visual pattern recognition and has been debated for its simplicity and its inability to compensate for unusual and complex visual displaces.

The Feature Models Theory is, in fact, a series of theories deriving from various geometric features that are used in experimentation. As stated by Pinker (1988):

> In these models, there are no templates for active shapes; rather, there are mini-templates or 'feature detectors' for simple geometric features such as vertical and horizontal lines, curves, angles, 'T-Junctions,' etc....The match between input and memory would consist of some comparison of the levels of activation of feature detectors in the input with the weights of the corresponding features in each of the stored shape representations, for example, the product of these two vectors, or the number of matching features. The shape that exhibits the highest degree of match to the input is the shape recognized. (pp. 6-7)

This theory, also, has serious drawbacks, nonetheless of which are its experimentation with simple geometric shapes, and that it does not take into account the relationship between the various feature detectors.

The Fourier Models Theory of image (pattern and shape) recognition, named after the trigonometry mathematician Fourier, proposes that the recognition of patterns and shapes of images depends on their trigonometric spatial analysis, their frequency of appearance, and the degree of their bright and dark intensity. As further explained by Pinker (1988):

> In long-term memory each shape would be stored in terms of it's Fourier transform. The Fourier transform of the image would be matched against the long-term memory transforms and the memory transform with the best fit to the image transform would specify the shape that is recognized. (p. 9)

The Structural Descriptions Theory proposes that the recognition of visual images (via patterns and shapes) is achieved by matching visual inputs to existing ones in the long-term memory, structurally or symbolically matching each part separately and by compensating for its relationship to the whole.

Critics of this theory argue that this is not really a full shape and pattern recognition theory and suggest that this theory: "by itself does not specify what types of entities and relations each of the units belonging to structural description corresponds to, or how the units are created in response to the appropriate patterns of retinal stimulation" (Pinker 1988, p. 12).

The Marr-Nishihara Theory (1978), named after the perceptual psychology scholars David Marr and Keith Nishihara, proposes that early visual processing culminates in the construction of representation called 2 1/2D sketch, designed by these two scholars and defined by Pinker (1988) as follows: "The 2 1/2D sketch is an array of cells, each cell dedicated to a particular line of sight from the viewer's vantage point....The 2 1/2D sketch is intended to gather in one presentation the richest information that early visual processes can deliver" (pp. 14-15).

What Marr and Nishihara sought to do was to explain two fundamental problems in all previous theories; first, that non of these theories specified precisely where perception ends and cognition begins and, second, that they did not pay attention to what, in general, the shape recognition process must do or what specific problems it is designed to solve. Actually, the Marr-Nashihara 2 1/2 sketch theory does two things: (a) it examines the nature of the recognition problem to separate early vision from recognition and from visual cognition in general, and (b) it provides an explicit theory of three-dimensional shape recognition that is built on such fundamentals.

Regardless of these advantages, however, even this theory has serious drawbacks underlined by Pinker (1988) as follows:

> The 2 1/2D sketch itself is an ill-suited to matching inputs against stored shape representations for several reasons. First, only the visible surfaces of shapes are represented...Second, the 2 1/2D sketch is view-point specific...Furthermore, objects and their parts are not explicitly demonstrated. (p. 17)

Finally, the Massively Parallel Models theory was developed by several perceptual psychologists as an alternative approach that provides different types of solutions to the issue of visual recognition. Among them, Attneave (1982), Hrechanyk and Bellard (1982, and Hinton (1981) have suggested the so-called Massively Parallel Models Theory is actually a model of shape recognition using massively parallel networks of simple interconnected units, rather than sequences of operations performed by a single powerful processor.

Assessing the validity and reliability of this theory, Pinker (1988) concludes that:

> In general, massively parallel models are effective at avoiding the search problems that accompany serial computational architectures. In effect, the models are intended to assess the goodness-of-fit between all the transformations of our input patterns and all the stored shape descriptions in parallel, finding the pair with he highest fit at the same time. (p. 36).

Several scholars of visual recognition, including Pinker (1988), concluded that these models are still underdeveloped and therefore it is difficult to evaluate their validity and reliability without further investigations and verifications.

The process of remembering or reasoning about shapes or objects that are not currently before us and must be retrieved from memory or constructed from a description is referred to as visual imagery, or thinking in pictures. Unlike visual recognition (which is a neurophysiological process turned into a mental activity), visual imagery is a mental process that may or may not become a physical activity. It is the mind that creates imaginary codes--images--which may or may not exist and which may or may not be recalled from the reservoirs of our memory banks. Usually written or aural descriptions of events, situations, objects, etc., assist the mind to create corresponding images, aided always by memory recalling one's experiences. However, in our sleep or in daydreaming we create images--usually unconventional and unusual actions of our imagination--subconsciously and involuntarily. For the average person consciousness or normal, and unconscious or unusual visual thinking is the result of one's own idiosyncratic nature; although the former can be taught or reinforced, the latter is uncontrollable.

Visual imagery, or thinking in pictures, is a mental process which is much more complicated and difficult to observe, to study, and to measure scientifically. For this reason (a) a great number of philosophical objectives and speculations regarding imagery have been made over the years and a substantial body of literature has been created and (b) several sound theories regarding visual thinking or visual imagery have been developed. The former is beyond the scope of this study, however, the latter will be briefly reviewed next.

The literature on this issue is immense, starting with the classic works of Arnheim's (1969) Visual Thinking and McKim's (1980) Experiments in Visual Thinking. It identifies five distinct theories herein discussed as points of view of those researchers, scientists, and cognitive psychology scholars who developed them.

The Zenon Pylyshyn (1981) point of view states, in effect, that imagery is not a distinct cognitive module but a representation of general semantic knowledge. It consists of the use of general thought processes to stimulate physical or perceptual events, based on tacit knowledge of how physical events unfold (Pinker, 1988).

The Allan Pairio (1971) point of view suggests that imagery uses representations and processes that are ordinarily dedicated to visual perception rather than abstract conceptual structures subserving thought in general. It proposes, additionally, that one of those representations used in perception and imagery has a spatial or array-like structure (Pinker, 1988).

The R. N. Shepard (1981) point of view proposes that in imagery a shape is represented by a two-dimensional manifold curved in three-dimensional space to form a closed surface, such as a sphere. Each position within the manifold corresponds to one orientation of the shape, with nearby positions corresponding to nearby orientations (Pinker, 1988).

The S. M. Kosslyn (1980) point of view claims that the medium, which he calls visual buffer (which can be substantiated with a computational model) is two-dimensional and Euclidean, and the position of the cells within the array corresponds to positions within the visual field. Kosslyn (1984) and his associates claim that once in the visual buffer, the pattern of activated cells can be rotated, scaled in size, or translated, and the resulting patterns can be examined by operations that detect shapes and spatial configurations (Pinker, 1988).

The G. E. Hinton (1979) point of view advocates that in visual imagery there are processes dedicated to the manipulation of spatial information, as it is also suggested by Kossylyn's (1980) model. It suggests that there is a spatial format for information represented in imagery that involves a global viewer-centered reference frame and there is an array-like scale within which the spatial disposition of the represented shape is specified (Pinker, 1988).

A representative theorist of visual cognition from the field of communication is Lester (1995), who has provided five such theories divided into two fundamental groups which he calls sensual theories, such as gestalt, constructivism, and ecological and perceptual theories such as semiotic and cognitive.

As Pinker (1988), who divides the process of visual cognition into representational and remembering, Lester (1995) also divides it into two stages which he calls sensational and perceptual; in reality he examines them as visual perception (stimulation, sensation, response) and visual cognition (biological and mental decodification of visual images). It is evident, from Lester's (1995) discussion and analysis of the five theories of visual communication, that he equates perception with cognition. However, as it is shown from the brief review of these theories that follows, he actually means cognition when he discusses perception.

The Gestalt Theory of Visual Recognition, developed by the German psychologist Max Wetheimer, states that all visual phenomena can be organized into various groups, which combined, create bigger units or configurations, which the brain receives, decodes, and recognizes. As Lester (1995) further explains:

> Gestalt psychologists further refined the initial work by Wetheimer to conclude that visual perception was a result of organizing sensual elements or forms into various groups. Discrete elements within a scene are combined and understood by the brain through a series of four laws of groupings: the law of similarity, the law of proximity, the law of configuration, and the law of common fate. (pp. 53-54)

The drawbacks of this theory are: (a) it is a visual perception theory rather than a visual cognition theory because it only describes how perception occurs and does not explain how the meaning of these images is given, (b) it is a stimulus-response explanation of the visual communication process, without ever referring to the importance of message and other factors involved in the recognition process, and (c) whereas it explains how the codification of visual images is manifested, it does not explain how the decodification, or understanding of visual images, is achieved.

The Constructivism Theory of Visual Recognition was developed by Julian Hockberg (1970), a perceptual psychologist who recognized the flaws of the gestalt theorists (mainly their inability to consider viewers' mental state during active viewing), and theorized that viewers' active participation results in the construction of gestalts (units or forms) with constant eye-fixations on scenes which the mind receives and combines into holistic structures or images (Lester 1995).

Among the major drawbacks of the constructivism theory of visual recognition are: (a) the theory does not explain how eye-fixations and experience interrelate to create the final picture, (b) the specific role played by memory during the recognition process is not clear, and (c) whereas the theory works with the use of simple figures and drawings, it is not clear how effective it will be with complex visual inputs such as fast moving images.

The Ecological Theory of Visual Recognition stems from the ecological theory of visual perception developed and published by Gibson (1979). The theory states, in effect, that the perception of objects in the environment is not determined by their size but by their scale or proportions, which remain constant during observation. On the bases of the number of surface grid units an object occupies in space, its size is said to be large or small. However, for Gibson (1979), neither the size nor the depth factor of objects require a high-level brain activity to be recognized because they are simple perceptual facts and direct perceptual experiences, which do not need extra mental calculations to be recognized. This distinction between ecological visual perception processes and cognitive activities in picture recognition is further explained by Lester (1995) as follows:

For Gibson perception is simply a matter of light striking objects and given the viewer enough information to determine whether the objects should be used or avoided...His ideas probably are best guess of how animals use visual perception, but humans learn to associate meanings with the objects they can see. Cognition is based on previous experiences, cultural factors, and linguistic abilities that contribute to the total concept visual perception. (pp. 60-61)

Like Gibson (1979), Lester (1995) falls into the same linguistic, or nomenclature, traditional error of equating visual perception with visual recognition, although in their discussion of these terms the difference becomes clear. This is one of the main drawbacks of the ecological theory of visual recognition. Another shortcoming is the inability of the theory to consider the viewers' state of mind and idiosyncratic nature during the visual recognition processes. Furthermore, this is purely a visual perception, not a visual recognition theory and, as such, it does not shed light on the issue of visual recognition other than helping to identify the two different processes involved in the study of visual communication, perception versus cognition.

<u>The Semiotics Theory of Visual Recognition</u> is one of the oldest theories of communication in general, which has been applied to all other areas and disciplines within the field, including visual perception and visual cognition. Lester (1995) classifies it as a perceptual theory of visual communication. However, he discusses it as cognitive theory recognizing that: "Although vision cannot happen without light illuminating, structuring, and sometimes creating perception, these two approaches [semiotics and cognition] stress that humans are unique in the animal kingdom because they assign complex meanings to the objects that they see" (p. 61).

In simple terms, the semiotic theory of visual recognition states, in effect, that every visual object is a sign that conveys a special meaning the viewer must learn to be able to decode (or connote) it properly. There are three types of signs, <u>iconic</u>, <u>indexical</u>, and <u>symbolic</u>, which not only determine the speed by which visual images are recognized, but also determine the degree of comprehension of the visuals, the highest of which is the symbolic, followed by the indexical, and the iconic. Explaining how this process takes place, Lester (1995) states:

> The study of signs is based on the idea that the hypocampus of the brain stores images in a symbolic form in order to recognize an object almost instantaneously. With instant identification of an image, either directly experienced or mediated, the brain can classify it immediately as helpful or harmful. (p. 63)

There is no doubt that the semiotics theory of visual recognition is both powerful and widely spread across the various academic disciplines. Yet it imposes certain obstacles to users as follows: (a) signs, or visual codes, must be learned to be readily decoded and recognized, (b) since each society creates its own signs, symbols, codes, etc. (deeply rooted in the culture), the denotation and connotation of visual images across cultures often are inaccurate, and (c) stemming from the linguistics theory, semiotics is a study of signs that carry pragmatic, syntactic, or semantic properties, from which only the syntactic signs are applicable to visual images. The various graphic elements that compose an image, such as lights, colors, and vectors are a collection of visuals that provide meaning for the viewer.

The Traditional Theory of Visual Recognition states that recognition and total understanding of visual images are the result of the dual process of biological (brain decodification activity) and mental (mind given meaning and reasoning). It goes beyond the definition given by perceptual psychologists (and even certain communication scholars) and becomes a meta-perceptual, or cognitive, rather activity in which the visual codes perceived are now organized and categorized by the brain's hypothalamus and they are processed to the mind, which involves memory, experiences, knowledge, and a host of other factors, to provide the appropriate connotations and meanings to the visual inputs.

Although Lester (1995), in his examination of the traditional visual cognition theory provides a lengthy discussion of the evolution of the theory, he still considers visual perception synonymous to recognition, stating that: "Visual perception is a function of the meaning we associate–through learned behavior or intelligent assumptions–with the object we see" (pp. 67-68). This becomes even more apparent when he considers, further, that the brain is a complex image processor. He suggests, along with various other cognitive psychologists, that either through alphabetizing visual images (theorized by such scholars as Biederman [1987] and Saint Martin [1990]), or through such mental activities as memory, projection, expectation, selectivity, habituation, salience, dissonance, culture, and words (theorized by Bloomer [1990]), the brain manages to translate the code of visual inputs into cohesive--holistic--forms, completed meaningful images. Neither perceptual psychologists nor communication researchers have managed to compare the biological activities of the brain with the mental processes of the mind. It is this particular factor that makes the traditional visual recognition theory more applicable and in par with recent advances in the fields of cognitive psychology and visual communication.

In summary, the clarification of the terms sensation, stimulation, perception, and cognition, and the review of the main theories of visual cognition and visual imagery provide the bases for the discussion of the model of visual recognition proposed in this study, diriving from the analysis of the various biological and mental activities of the brain.

From Visual Precepts to Visual Concepts

A great number of electromagnetic, chemical, and generally biological (physical) and mental (psychological) activities are involved in visual perception. The instantaneous manner and speed by which they occur have made it difficult in the past to isolate the steps, to observe them closely, to study them thoroughly, and to describe them accurately. Recently, however, due to enormous technological advancements in measuring devices and computer assisted research techniques many of these complex activities have been observed and studied by neurophysiologists and perceptual psychologists. The student of visual communication media, and generally the constructors of visual images, must acquire the basic knowledge regarding these complex activities, which will help them to produce better, more effective, and more appropriate pictures and programs. This is precisely the objective of this section, and indeed the purpose of this paper: to unveil the processes and activities involved not only during visual stimulation, visual reception (or perception), and visual codification, but, foremost, the various steps during which the visual precepts, codified visual units, become visual concepts, holistic visual images.

In previous studies I have provided the three main steps involved in the perceptual process of images, namely distal (stimulation), proximal (perception), and perceived (recognition), (Metallinos, 1996).

In the model of perception and cognition, I have identified the environmental (sensations), the electromagnetic (perceptions), and the electrodermal (concepts) that occur during the process of visual communication (Metallinos, 1996).

Whereas the description and discussion of the first two processes—environmental sensations and electromagnetic perception—are easily understood and readily accepted, the third process, the electrodermal concept, the final recognition of visual images, is not. In other words, the stimulation and perceptual activities involved in visual communication, which lead to the codification of the visual data, can be observed and studied. But the decodification and final interpretation of this data by the brain is more complex, difficult to observe, and almost impossible to locate precisely and accurately. It requires the close study and analysis of (a) the physiology of the brain, (b) the functions performed by the brain, and (c) the role played by a great number of psychological or mental factors (such as culture, habituation, salience, projection, expectations, selectivity, dissonance, words, and memory, explained by Bloomer [1990]) and grouping (psychological closure), texture (or contrast and figure–ground distinctions), time (or duration) and timing, motion (or directional vectors), depth (or three-dimensionally), and imagination or enlargement of memorable experiences), which I have discussed in previous studies (Metallinos, 1996). Among all these factors the most fundamental is memory, not only because memory is closely related to all these mental activities, but above all because memory is the basic ingredient for the transformation of the biological precepts (visual and auditory codes) into mental concepts (recognized images).

As the authors mentioned above, along with a great number of others, have provided thorough and detailed discussions on each of these mental functions, including the anatomy of the brain, I will concentrate herein on the role of memory--the most significant mental factor--in the brain's transformation of precepts to concepts; that is, how is visual and auditory information is received, stored, and retrieved?

The process of visual and auditory perceptions, that is how we receive information, has been repeatedly discussed. How such information is stored and how it is remembered are the activities that involve the role of memory in transforming visual and auditory codes (precepts) into completed entities (concepts) discussed next.

Neuroscientists such as Bloom, Lazerson and Hofstadter (1985) and Pines (1986) have identified and studied various categories of memories based on their duration and function. On the basis of the length of time that memories are stored, Bloom et al. (1985) indicate that there are three phases: (a) the so-called <u>immediate memory</u>, the extremely short-term memory in which information pieces are stored for only a few seconds and which may or may not transfer to another phase of memory, (b) the <u>short-term memory</u>, where information inputs are held for several minutes, and (c) <u>long-term memory</u>, where storage of information may last for hours or for a lifetime.

On the basis of the functions that memories perform they seem to be either <u>procedural</u> or <u>declarative</u>, two different systems each located in different parts of the human brain, which Pines (1986) explains:

Procedural memory, a memory for skills, probably develops earlier in life than declarative memory, the ability to recall facts. Fact memory appears to center in the hippocampus, amigdala, and part of the thalamus. Procedural memory, believed more widely dispersed, is therefore less subjective to impairment by illness or injury. (p. 359)

The procedural and declarative memory systems were the basis for the development of two distinctive schools of learning in the field of psychology. The behavioral, which advocates that people memorize and learn skills and habits by reinforcement and conditioning, and the cognitive school of thought advocating that people memorize and acquire information and knowledge intentionally by their own free will without expecting to be rewarded. This type of learning is a function of the higher brain and conscious mind (Pines, 1986). Behavioral learning, stemming from the procedural memory system, provides knowledge as how to do things, whereas cognitive learning, stemming from the declarative memory system, provides accurate records of particular experiences and a sense of familiarity about these experiences (Bloom et al., 1985). In fact, certain experiments have indicated that procedural memory, which generates behavioral learning, occurs as a biochemical or biophysical activity, which occurs only in the neural circuit directly involved in procedural learning. On the other hand, declarative memory, which generates cognitive learning, occurs as an activity of constant remodeling of neural circuitry and seems to be a psychological process or mental activity (Bloom et al, 1985).

During our moving image recognition both procedural and declarative memories are in operation in order to obtain the necessary ingredients to match the incoming information. However, the degree of development of one memory system over the other is analogous to the individual's own preference. The length of the information storage and the procedural or declarative systems we choose to maintain our memories in are connected to a series of other brain and mind operations, the subtotal of which constitutes the idiosyncrasy of a particular individual. As Pines (1986) suggests: "What we choose to store in our long-term memory is closely tied to our emotions" (p. 369). Consequently, an individual's stored experiences are subjective and, as such, his/her retrievals of these experiences are also idiosyncratic. It is this indisputable factor that makes the task of the visual communication media producer/director harder because the constructed visual messages must be readily received and codified (as precepts) so that individual viewers can recognize them with the same degree of readiness and accuracy as concepts asserted by the reservoir of their memories and knowledge. Therefore, in the visual communication processes both the constructors of moving images (senders) and their consumers (receivers) must have a reciprocal understanding of the codification–decodification processes of the brain and the mind respectively. The more one knows, retains, uses, or practices the visual communication processes and activities, the more one understands how the codes are structured and how they should be interpreted.

Given that (a) the observations of the transformation of biological precepts to mental concepts in the recognition of moving images cannot be seen but only inferred, (b) the memory storage and retrieval systems of learning are highly idiosyncratic, and (c) memory, as a mental activity is closely related to a host of other brain and mind activities mentioned above, a concise, scientifically sound, model of visual recognition is impossible. It can only be inferred and schematically given so that the major steps and subsequent activities that take place in these steps can be illustrated and oversimplified. One such model, in addition to the ones I have previously mentioned, was given by Frisby (1980). This is not really a visual recognition model but a schematic illustration of how the human visual system works. However, it suggests what is going on in the last two steps in the process that transforms the biological precepts to mental concepts, which Frisby (1980), called segmentation (step #6) and object recognition (step #7). The close resemblance to my model of stimulation, perception, and recognition is apparent, particularly since Frisby's steps 1-5 (scene retinal image, gray-level description, lightness and brightness computation, and feature descriptions) constitute my stimulation–perception steps, and his steps 6 and 7 (segmentation and object recognition) resemble my recognition step, which consists of the codification-decodification activities of the brain and the mind.

In summary, the transition of the visual image input from visual codes (precepts) to visual images (concepts) is the result of combination of biological and mental activities, with the support of memory that acts as the catalyst for this transformation.

We turn now to the last section of this inquiry with suggestions for the construction of artistic moving images.

Construction and Artistic Synthesis of Moving Images

Now that we have seen how the brain and the mind work to transform visual and auditory precepts into holistic concepts, we are able to provide a series of guidelines and suggestions to the constructors of moving images--primarily television images--regarding the artistic synthesis of such images based on the traditional theories of visual recognition and visual imagery.

The artistic synthesis of televised images, as well as all moving images, depends on four major factors deriving from the particular instruments, materials, and techniques that comprise the visual communication medium. In the case of film and television media such factors are (a) light and color, (b) framing or staging, (c) audio setting and sound selection, and (d) editing or sequencing of the visual and auditory elements. Each of these factors, as visual communication media components, has been studied as a unique entity and contributory factor in the total synthesis of film and television pictures. As constructors of moving images, via the visual communication media of film and television, we must understand how each of these components works, how the instruments of each medium must be handled, and what techniques are more appropriate for the lighting, framing, audio, and editing of the particular program (Metallinos, 1996; Wurtzel, 1983).

What we learned from the theories of visual and auditory perception and primarily the theories of visual recognition, is that we are not totally free in composing pictures, primarily pictures with an artistic or aesthetic merit. We must follow the composition guidelines deriving from the traditional fine arts, enhanced with new scientific information regarding the visual recognition processes by the brain and the mind. The four main suggestions regarding (a) lighting and color, (b) framing and camera angle, (c) audio and sound effects, and (d) sequencing, or editing visuals and sounds in television pictures are provided herein.

1. Light the set for bright and dark intensity.

   According to Fourier models theory discussed earlier, image recognition depends, among other things, on the degree of their bright and dark intensity. It is, therefore, important to consider lighting the set or the objects to be videotaped not only for mood and atmosphere purposes, but for maximum visibility, overall picture clarity, and inevitably easy recognition. This is a rule that has not always been considered by producers/directors of television pictures, who often abuse this fundamental rule with perceptual gimmicks.

2. Frame the scenes for maximum symbolic and structural recognition.

   According to the structural description theory of visual recognition of patterns and shapes, the visual elements retained the most are the new ones which are stored in the long term memory and are constructed either as symbolic representations of the objects depicted, or as propositions, which correspond to the various parts of the objects. Consequently, constructors of television images should pay special attention to those images that provide better clues for recognition and warrantee viewers' interests, attention, and final appreciation of visual pictures.

3. Coincide the program's sounds with the images and vice versa.

   Matching pictures with sounds or applying appropriate sounds to particular visuals is equal to thinking in pictures. Because sounds remind us of images, so do images bring to mind certain corresponding sounds. The imagery theories that explain how we create images in our minds must be employed in the case of selecting the audio in television production. For example, the use of general thought processing to stimulate physical or perceptual events, based on tacit knowledge of how physical events unfold, developed by Pylyshyn (1981), and refering to imagery, is equally applicable to sound. The constructors of visual images must be aware of these principles, should employ them, and should never ignore them.

4. Edit the sequences analytically or associatively.

   All visual recognition theories reviewed earlier suggest that successful recognition of visual images depends on the ways by which the images succeed each other. Such factors as spatial analysis, frequency of appearance, spatial relationships, visual unit interconnections, are means of sequencing that maximize the recognition of picture sequencing. This, in turn, coincides with the rules of composition in television (and film) editing divided into continuity or complexity editing strategies, the former employing a diagnostic form of editing and the latter employing a thematic one (Metallinos, 1996). It is mostly in this area that the visual communication media rules of composition directly coincide with the scientific rules of scene sequencing, which visual communication media producer/directors should be familiar with and they should not ignore.

In summary, the constructors of visual communication media programs should enhance their knowledge of picture composition, combining the communication media related rules of composition with the scientific theories of visual image recognition and imagery. Only then will visual communication media such as television be seen and considered a medium capable of producing aesthetically pleasing pictures and, subsequently, artistic television programs.

References

Arnheim, R. (1969). Visual thinking. Berkeley, CA: University of California Press.

Attneave, F. (1982). Pragnayz and soap bubble systems: A theoretical exploration. In J. Beck (Ed.), Organization and representation in perception (pp. 102-108). Hillsdale, NJ: Lawrence Erlbaum.

Behnke, R. R. (1970). Psychological technologies. In P. Emmert & W. D. Brooks (Eds.), Methods of research in communication (pp. 429-452). Boston, MA: Houghton, Mifflin.

Biederman, I. (1987). Recognition by components: A theory of human image understanding. Psychological Review, 94, 115-147.

Bloom, F. E., Lazerson, A., & Hofstadter, L. (1985). Brain, mind, and behavior. New York: W. H. Freeman.

Bloomer, C. M. (1990). Principles of visual perception. New York: Design Press.

Frisby, J. P. (1980). Seeing: Illusion, brain, and mind. Oxford, England: Oxford University Press.

Gibson, J. J. (1979). The ecological approach to visual perception. Boston: Houghton–Mifflin.

Haber, R. N. (Ed.). (1968). Contemporary theory and research in visual perception. New York: Holt, Rinehart, & Winston.

Hinton, G. E. (1979). Some demonstrations of the effects of structural descriptions in mental imagery. Cognitive Science, 3, 231-250.

Hinton, G. E. (1981). A parallel computation that assigns canonical object-based frames of reference (pp. 209-225). Proceedings of the International Joint Conference on Artificial Intelligence. Vancouver, Canada.

Hockberg, J. (1970). Attention, organization, and consciousness. In D. Mostosfky (Ed.), Attention: Contemporary theory and analysis (pp. 210-223). New York:

Appleton-Century-Crofts.

Hrechanyk, L. M., & Bellard, D. H. (1982). A connectionist model of form perception. Proceedings of the IEEE Special Workshop on Computer Vision (pp. 74-92). Ringe, NH.

Kosslyn, S. M. (1980). Image and mind. Cambridge, MA: Harvard University Press.

Kosslyn, S. M., Brunn, J., Cave, K., & Wallash, R. (1984). Individual differences in mental imagery ability: A computational analysis. Cognition, 18, 195-243.

Lester, P. M. (1995). Visual communication: Images with messages. Belmont, CA: Wadsworth.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. Proc. R. Soc. Lond, 200, 269-294.

Martin, I., & Venables, P. H. (Eds.). (1980).The technique in psychophysiology. New York: John Wiley & Sons.

McKim, R. H. (1980). Experiments in visual thinking (2nd ed.). Belmont, CA: Wadsworth.

Metallinos, N. (1983). Biometric research instruments. In R. A. Braden & A. D. Walker (Eds.), Seeing ourselves: Visualization in a social context: Readings of the 14th annual conference of the International Visual Literacy Association (pp. 171-175). Blacksburg, VA: The International Visual Literacy Association.

Metallinos, N. (1996). Television aesthetics: Perceptual, cognitive, and compositional bases. Mahwah, NJ: Lawrence Erlbaum.

Murch, G. M. (1973). Visual and auditory perception. New York: Bobbs-Merrill.

Pairio, A. (1971). Imagery and verbal processes. New York: Holt, Rinehart, & Winston.

Pines, M. (1986). Landscapes of the mind. In R. M. Poole (Ed.), The incredible machine (pp. 325-377). Washington, DC: National Geographic Society.

Pinker, S. (Ed.). (1988). Visual cognition. Cambridge, MA: The MIT Press.

Pylyshyn, Z. W. (1981). The images debate: Analog media versus tacit knowledge. Psychology Review, 88, 16-45.

Saint-Martin, F. (1990). Semiotics of visual language. Bloomington, IN: Indiana University Press.

Shepard, R. N. (1981). Psychophysical complementarity. In M. Kubory & J. Pemerantz (Eds.), Perceptual organization. Hillsdale, NJ: Lawrence Erlbaum.

Springer, S. P., & Deutch, G. (1985). Left brain, right brain (rev. ed.). New York: W. H. Freeman.

Tarroni, E. (1979). The aesthetics of television. In H. Newcomb (Ed.), Television: The critical view (2nd. ed.), (pp. 437-461). New York: Oxford University Press.

Wurtzel, A. (1983). Television production (2nd. ed.). New York: McGraw–Hill.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: THE TRANSFORMATION OF THE BIOLOGICAL PERCEPTS INTO MENTAL CONCEPTS IN RECOGNIZING VISUAL IMAGES

Author(s): Dr. Nikos METALLINOS

Corporate Source: CONCORDIA UNIVERSITY, MONTREAL, CANADA

Publication Date: Janu 1998

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 [X]	Level 2A [ ]	Level 2B [ ]
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: Nikos Metallinos

Printed Name/Position/Title: Nikos METALLINOS, PROFESSOR

Organization/Address: CONCORDIA UNIVERSITY, MONTREAL, CANADA

Telephone: (514) 848 2536

FAX: (514) 848 4513

E-Mail Address: 105757.2731 @ compuserve.com

Date: Janu 1998

(over)