30. Taddei-Ferretti, C., C. Musio & R.F. Colucci, (1994), Top-down intrerference in visual perception, *ICANN '94*, M. Marinaro & P.G. Morasso, eds, Berlin, Heidelberg, New York, Springer, 30-33.

31. Taddei-Ferretti, C., C. Musio & S. Santillo, (1995), Non linearity of the interactions between bottom-up and top-down signals in multistable visual perception, *Il Nuovo Cimento D*, 17:941-947.

32. Taddei-Ferretti, C., C. Musio, R.F. Colucci & M. Martino, (1993), Multistable ambiguous pattern perception, *WCNN-INNS '93*. vol. II, Lawrence Erlbaum Ass., Hillsdale, N.J., vol. II, 155-158.

33. Ullmann, S., (1984), Visual routines, *Cognition*, 18:97-159.

Toward a Theory of Visual Abductive Thinking

Lorenzo Magnani
Guido Previde Massara
Computational Philosophy Laboratory
Department of Philosophy
University of Pavia
I-27100 Pavia – Italy
lorenzo@philos.unipv.it, guido@philos.unipv.it

Abstract

The general objective is to consider how the use of visual mental imagery in thinking may be relevant to hypothesis generation. There has been little research into the possibility of visual and imagery representations of hypotheses, despite abundant reports that imaging is crucial to scientific discovery. Glasgow and Papadias have suggested an interesting technique for combining image–like representations and processing with linguistic information. The system they describe can make inferences using such cognitive representations as visual mental imagery ones: it achieves conclusions without laborious chains of inferences. We plan to explore whether this kind of hybrid imagery/linguistic representation can be modified and used to model image–based hypothesis generation, that is to delineate the first cognitive and computational features of what we call *visual abduction*.

1 Introduction

Many psychological and physiological studies have been carried out to describe the multiple functions of mental imagery processes: there exists a visual memory that is superior in recall; humans typically use mental imagery for spatial reasoning [1]; images can be rebuilt in creative ways [3]; they preserve the spatial relationships, relative sizes, and relative distances of real physical objects [9]; for a more complete list, see Tye [28].

Kosslyn introduces visual cognition as follows:

> Many people report that they often think by visualizing objects and events [...] we will explore the nature of visual cognition, which is the use of visual mental imagery in thinking. Visual mental imagery is accompanied by the experience of seeing, even though the object or event is not actually being viewed. To get an idea of what we mean by visual mental imagery, try to answer the following questions: [...] How many windows are there in your living room? If an uppercase version of the letter *n* were rotated 90° clockwise, would it be another letter?" ([11] p. 128).

We can build visual images on the basis of visual memories but we can also use the recalled visual image to form a new image we have never actually seen. Certainly, imagery is used in everyday life, as illustrated by the previous simple answers, nevertheless imagery has to be considered as a major medium of thought, as a mechanism of thinking relevant to hypothesis generation. Some hypotheses naturally take a pictorial form: the hypothesis that the earth has a molten core might be better represented by a picture that shows solid material surrounding the core. There has been little research on the possibility of visual imagery representations of hypotheses, despite abundant reports (e.g. Einstein and Faraday) that imaging is crucial to scientific discovery. It is well–known that the German chemist Kekulé used spontaneous imagery to discover the structure of benzene; Watson and Crick have reported the use of mental imagery in the interpretation of diffraction data and in the determination of the structure of the DNA molecule [7,13,15,16,26,27].

2 Computational Scheme and Abduction

The central theme of the recent imagery debate in the cognitive science has concerned the problem of representation.

How can we represent images? Are mental images represented depictively in a picture or like sentences of descriptions in a syntactic language?

According to Kosslyn's *depictionist* or *pictorialist* view [10], mental images are *quasi–pictures* represented in a specific medium called the *visual buffer* in the mind. Kosslyn builds the analogy that visual information stored in memory can be displayed as an image on a CRT screen. Kosslyn's model of mental imagery proposes three classes of processes that manage images in the visual buffer: the *generation* process forms an image exploiting visual information stored in long–term memory, the *transformation* process (for example, rotation, translation, reduction in size, etc.) modifies the depictive image or views it from different perspectives, the *inspection* process explores patterns of cells to retrieve information such as shape and spatial configuration. According to Pylyshyn's *descriptionist* view [21] mental imagery can be explained by the tacit knowledge used by humans when they simulate events rather than by a pictorialist view related to the presence of a distinctive mental image processor.

As shown by certain experimental results, Hinton proposes that imagery involves viewer–centered information appended to object–centered structural descriptions of images [6]. Finke [2] proposes five unifying principles of mental imagery to summarize some of the properties of imagery that a computational model should represent. Finally, Tye has recently suggested a theory, related to some of the aspects of Marr and Nishihara's theory of visual perception [14], in which mental images are represented as interpreted, symbol–filled arrays [28].

In Glasgow and Papadias' computational model [4,5] the imagery processes are simplified by a depictive knowledge representation scheme that involves inferencing techniques related to generation, inspection and transformation of the representation. According to Kosslyn's cognitive model, the knowledge representation scheme of mental imagery is composed of two different levels of reasoning, *visual* and *spatial*, the former concerned with what an image looks like, the latter depending on where an object is located relative to other objects. The different representations of these ways of reasoning exist at the level of working–memory and are generated (according to Kosslyn's theory of imagery [9]) from a *descriptive* representation of an image stored in long–term memory in a hierarchical organization. The separate long–term memory contains visual and spatial information as descriptions.

The spatial representation outlines the image constituents in a multi–dimensional, symbolic array, that maintains spatial and topological properties and avoids distracting details, in accordance with the features of the image domain and the questions involved;

				Sweden	
Scotland / Wales England			Denmark		
		Holland	Germany	Germany	
		Belgium			
		France	Italy	Yugoslavia	Yugoslavia
Portugal	Spain		Italy		Greece

Embedded symbolic array representation

Figure 1: Embedded symbolic array representation [5]

moreover, the arrays are *nested*, to express multiple levels of the structural hierarchy of images (Fig. 1).

The spatial representation is processed using primitive functions that *transform* and *inspect* the symbolic array. The functions implemented express simple operations and reasoning techniques based on the transformation and inspection of image representations.

Moreover, a multi–dimensional array is able to depict spatial (e.g. *left–of, north–of*) and topological (e.g. *adjacent–to, inside*) relations among image parts and can also denote non–spatial dimensions such as speed. Sometimes, a part occupies more than one element in the array, to capture all the spatial relationships involved. Hence, an image frame is related to an individual image or to a class of images sharing the same features. Images can be represented in a viewer– or an object–centered frame of reference: unlike Kosslyn's approach, in Glasgow and Papadias' computational model [4] image representations are typically three–dimensional and object–centered (viewer–independent).

The "uninterpreted" visual representation depicts the space occupied by an image as an *occupancy array*. The related primitive functions concern operations for manipulating and retrieving geometric information, such as *volume*, *shape* and *relative distance* (rotate, translate and zoom, which change the orientation, location or size of a visual image representation). In a nested symbolic array each symbol in the array corresponds to a frame in long–term memory: we can extract a subimage and express it by its symbolic array parts.

The main objective is now to consider how the use of visual mental imagery in thinking may be relevant to hypotheses generation. What is abduction? After having illustrated the concept of abduction we propose an image–based hypothesis generation.

Abduction is becoming an increasingly popular term in AI [18,19,20,23,24,25] especially in the field of medical KBSs [8,12,22]. A hundred years ago, Peirce [17] interpreted abduction essentially as an inferential *creative process* of generating a new hypothesis.

There are two main epistemological meanings of the word abduction: 1) abduction that only generates plausible hypotheses (*selective* or *creative*)[1] and 2) abduction considered as *inference to the best explanation*, that also evaluates hypotheses. All we can expect of our "selective" abduction, is that it tends to produce hypotheses that have some chance of turning out to be the best explanation. Selective abduction will always produce hypotheses that give at least a partial explanation and therefore have a small amount of initial plausibility. In this respect abduction is more efficacious than the blind generation of hypotheses.

We should remember, as Peirce noted, that abduction plays a role even in relatively simple *visual phenomena*. *Visual abduction*, a special form of abduction, occurs when hypotheses are instantly derived from a stored series of previous similar experiences. It covers a mental procedure that tapers into a non–inferential one, and falls into the category called "perception". Philosophically, *perception* is viewed by Peirce as a fast and uncontrolled knowledge–production procedure. Perception, in fact, is a vehicle for the instantaneous retrieval of knowledge that was previously structured in our mind through inferential processes. By perception, knowledge constructions are so instantly reorganized that they become habitual and diffuse and do not need any further testing. Many visual stimuli are ambiguous, yet people are adept at imposing order on them: "We readily form such hypotheses as that an obscurely seen face belongs to a friend of ours, because we can thereby explain what has been observed" ([24] p. 53). This kind of image–based hypothesis formation can be considered as a form of *visual abduction*.

3 Visual Abduction

3.1 Image–Based Explanation

We will now discuss, from a computational perspective, a *visual abductive problem solving strategy*. To provide manageable bounds to our very general objective, i.e. to analyze the role of visual hypothesis generation, which is so crucial to scientific discovery, we have initially limited ourselves to the subtask of illustrating some structurally similar examples from the field of common sense reasoning, where it is very easy to find many cases dealing with what we have just called visual abductive problem solving. Moreover, we have limited ourselves to the spatial representation.

The spatial representation does not add information that cannot be expressed by propositions; notwithstanding this, the spatial representation is not computationally equivalent to a descriptive one. In several imagery–related tasks (e.g. inspecting) spatial representation may reduce the computational complexity of the solution.

Although there is considerable agreement concerning the existence of a high–level visual and spatial medium of thought as a mechanism relevant to abductive (selective and creative) hypothesis generation, the underlying cognitive processes involved are still not well understood. Notwithstanding this, we will attempt to work around this gap in our understanding: although describing a computational model able to "imitate" the real ways the human brain works when it makes visual abductions would be best, our

[1] To illustrate from the field of medical knowledge, the discovery of a new disease and the manifestations it causes can be considered as the result of a creative abductive inference. Therefore, *creative* abduction deals with the whole field of the growth of scientific knowledge. This is irrelevant in medical diagnosis where instead the task is to *select* from an encyclopedia of pre–stored diagnostic entities [22].

primary concern is its expressiveness, efficiency and inferential adequacy, rather than its explanatory and predictive power as regards psychological research.

Let us consider the example above concerning the ambiguity of an obscurely seen face: more generally, we can face an *initial* (eventually)[2] observed image in which we recognize a problem to solve. For example, given a visual or imagery datum, we may have to explain: 1) the absence of an object; 2) why an object is in a particular position; or 3) how an object can achieve a given task moving itself and/or interacting with the remaining objects in the scene/image. How can "visual" reasoning perform this explanation? To answer this question it is necessary to show how visual abduction may be relevant to hypothesis generation, that is, how an *image–based explanation* is able to solve the problem given in the initial image.

Faced with the initial image, in which we have previously recognized a problem to solve, as stated above, we have to work out an *imagery hypothesis* that can explain the problem–data. Thus, the formed image will acquire a hypothetical status in the inferential abductive process at hand.

The generation of a "new" imagery hypothesis can be considered the result of the *creative abductive* inference previously described; in this respect we can consider how the imagery representations of new hypotheses lead to scientific discovery. The selection of an imagery hypothesis from a set of pre–enumerated imagery hypotheses, stored in long–term memory, also involves abductive steps, but its creativity is much weaker: this type of visual abduction can be called *selective* [12] (see Sect. 2).

3.2 Imagery Hypotheses

As stated above, to illustrate the role of visual hypotheses generation we have initially limited ourselves to the subtask of illustrating some structurally similar examples from the field of common sense reasoning. It is not difficult to mention a few cases in which we recognize a visual or imagery datum as containing a problem to solve. As we have seen in the previous section, we could have to explain: 1) the absence of an object; 2) why an object is in a particular position or 3) how an object can achieve a given task moving itself and/or interacting with the remaining objects of the scene/image.

We shall clarify the first case with the following example, which is based on common sense reasoning: we see a broken horizontal glass on the floor, near the table. On the floor there are also some leaves and we see that the window is open. If we retrieve from long–term memory another *visual* (imagery) description still containing the glass (intact), the table, and the window, and we recognize this new representation as being a slightly different version of the previous one, we have to explain the presence of the leaves and broken glass in the initial image. They constitute an *anomaly* that needs to be solved (explained). If we can link the leaves to the presence, say, of wind, we are in turn transported to a new imagery explanatory hypothesis.

The second case deals with the capacity to justify the absence of a given object in a scene. Let us consider the following example: one of our friends is accustomed to travel the same route every day. The road passes near to a little bridge, under which ducks can usually be *seen* swimming. On a particularly cold day our friend does not *see* the ducks. He asks himself where the ducks could be, but, since he has never seen any ducks in a different setting, while he is able to detect the anomaly he is unable to explain it. The

[2] Of course the initial spatial image can be the representation of a real case.

imagery explanatory reasoning is impossible: therefore, it is stopped. On the contrary, if our friend had previously seen the ducks, say, under the roof of a farmhouse, once he has detected the absence of the ducks he can retrieve from long-term memory the image of the ducks sleeping under the roof. The imagery explanatory hypothesis is immediately achieved.

The third case deals with the well-known monkey-banana problem. In a room there is a banana, a box, and a monkey. The monkey cannot reach the banana because it is on the ceiling, but it can push the box to a point below the banana, climb on top of it and so reach the banana. Every visual representation of the effect of a sequence of actions the monkey can perform may be considered as an hypothesis generation. Such an hypothesis, if successful, is viewed as the one that gives a solution of the problem.

3.3 Visual Abduction System [VASt] Structure

We are developing a system (VASt) that focuses on spatial reasoning tasks and is ontologically organized into a *universe* of nested *spatial worlds.* At the computational level, the universe is composed of a collection of *array representations* that describe each spatial world, and of a *navigation device.* Thus, the different worlds can communicate with each other by means of the navigation device. Indeed, each spatial world is represented, according to Glasgow and Papadias' computational model, by a nested symbolic array and by a collection of rules identifying it (e.g. Newtonian mechanics). Imagery objects [3] are of two kinds: *active* and *passive.* Both kinds of objects "know" the rules of the world where they are included; moreover, they are individually characterized by a family of functions that imply some definite kinds of action they are able to perform. Finally, 1) each active object may interact with the remaining objects in a world; and 2) each passive object in a world may only act after interacting with an active object.

To represent interesting cognitive relations explicitly among the objects of the universe of spatial worlds we have to enrich Glasgow and Papadias' knowledge representation scheme by means of topological and combinatorial considerations upon the hierarchy of spatial worlds. Since the universe of spatial worlds can be mathematically represented as a tree, we can consider its adjacency matrix. A well known theorem in graph theory allows us to manage easily paths of any length and therefore to consider a lot of relations beyond the mere inclusion (e.g. "to be next to but not included", "o be elsewhere but not far from here" and so on). Such a capability induces the possibility to match spatial worlds consistently with these new relations, in order to provide a detailed computational problem solving strategy for solving visual abduction problems described above (that is the explanation of the absence or of the position of an object).

We are developing a set of functions with several capabilities, such as object manipulation, array retrieval, mapping and so on. The principal interest is to design functions of cognitive interest in array theory. The complete set of such functions covers the functionality of the *navigation device.* Here are some examples. Consider a hierarchy of nested arrays (Fig. 2). We are trying to give a precise meaning to "cognitive" sentences like "Array A is more interesting than array C", "There is a stronger link between G and F than between G and I" or "The object represented by array A is better known that any other". Someone could wonder why we think array A is more interesting than C, seeing

[3]By object in this situation we mean nothing else that a cell in the array, to which corresponds in turn another array at the lower level of representation

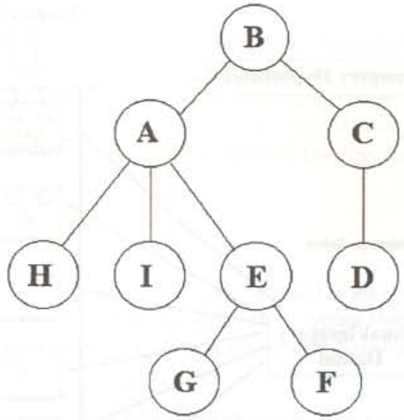

Figure 2: Hierarchy of nested arrays

that we do not know what the arrays A and C are representing. We think that such considerations can be simply suggested by the topological structure of the hierarchy of the arrays (that is, actually, a tree).

The basic idea is to take into account not only paths of length one between the nodes (e.g. the path E–F in the above illustration), but also paths of length more than one (e.g. the two-lenght path G–E–F). Obviously, from a cognitive point of view, a path of length one is more "important" than a path of length two. In order to consider this situation we give different weights to n-length paths (for instance a n-length path might have a weight of $1/n!$ – one over n factorial –). So the path between G and F will be weighted 0.5 while the path between G and I will be weighted roughly 0.166 giving a justification of one of the above sentences. The next step is to answer a (seemingly) meaningless question: "How many paths are there between A and A ?". The answer is: "There are four paths of length two, seven paths of length four and so on". We claim that the sum of the weights of such paths could be a measure of the cognitive interest of the array. This allows us to give a meaning to sentences like "Array A is the most important".

The cognitive meaning of a two length path (namely A–I–A) could be represented with the following sequence of operations: retrieval of the image A – focusing of a particular (namely I) – return to the image A. The topological structure of the representation is clearly subjective, reflecting one's points of view. For instance I know my home much better than my neighbour's one, correspondingly the array representation of my home is far richer.

To solve the third visual abduction problem introduced above (monkey-banana problem) it is easy to use a classical means-ends analysis strategy. The room is represented by an array where all the objects are indicated by symbols on the array. If we describe a primitive distance function, say $dist(x, y)$, we are immediately transported to a new formulation of the goal, that is, dist(monkey,banana) = 0. Let us suppose the "object" monkey is supplied with a set of primitive functions, such as ***walk-right***, ***walk-left***, ***push-left***, ***push-right***, ***climb***, ***grasp***. It is possible to describe visually the effect of every function in the array. A feasible procedure is 1) choose an action 2) if this action leads to a decrease in the distance then 2.1) do it 3) else go to 1). Such an algorithm clearly does not always lead to the goal. It has been improved with a device that is capable of detecting station-

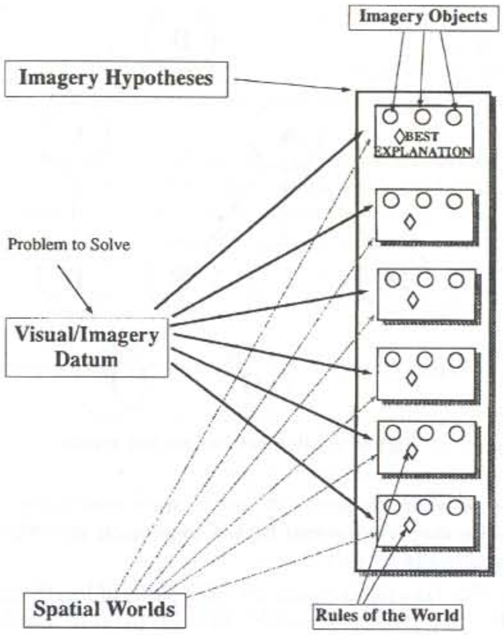

Figure 3: Visual abduction as image-based explanation

ary states which are not the solution of the problem. We can also avoid this impasse by adding to the distance function a penalty function which evaluates the "distance" of the actual state from the set of stationary states.

If we consider each new step leading to the goal (i.e. each new configuration of the array generated by the new positions of the objects) as an *imagery state*, we can say that the "monkey" (or "people" faced with the monkey–banana problem) forms different *imagery hypotheses* devoted to achieving the task. Thus, each step represents a particular imagery world. The configuration of the array in which the goal is achieved performs the image–based best explanation, i.e. the visual abduction: this generated abductive imagery hypothesis is the best explanation of the problem–data, and hence able to perform the planning task (Fig. 3).

Finally, if our spatial world represents a room, it can be supplied with a collection of rules of varying detail (e.g. Newtonian mechanics), yet it can also be supplied with completely different rules. Of course, using less rules results in the objects of the world being less constrained. For example, weakening the rules of Newtonian mechanics can lead to a new kind of spatial world, which is very abstract and considered as being a "virtual" one.

4 Conclusion

This paper has examined a kind of hybrid imagery–linguistic representation used to model image–based hypothesis generation (*visual abduction*): faced to an initial image, in which we have recognized a problem to solve, we have to generate an imagery hypothesis able

to explain problem-data.

Although describing a computational model able to "imitate" the real ways the human brain works when it makes visual abductions would be best, our primary concern has been expressiveness, efficiency and inferential adequacy, rather than its explanatory and predictive power as regards psychological research. Notwithstanding this, our system can contribute to discussions in the psychological area of the imagery debate.

References

[1] M.J. Farah (1988), The neuropsychology of mental imagery: converging evidence from brain-damaged and normal subjects. In J. Stiles-Davis, M. Kritchevsky & U. Bellugi (Eds.), *Spatial Cognition - Brain Bases and Development*, Hillsdale, NJ, Erlbaum, 33-56.

[2] R.A. Finke (1989), *Principles of Mental Imagery*, Cambridge, MA, The MIT Press.

[3] R.A. Finke & K. Slayton (1988), Explorations of creative visual synthesis in mental imagery, *Memory & Cognition, 16,* 252-257.

[4] I.J. Glasgow (1993), The Imagery Debate Revisited: a Computational Perspective, *Computational Intelligence, 9(4),* 309-333.

[5] I.J. Glasgow & D. Papadias (1992), Computational imagery, *Cognitive Science, 16(3),* 355-394.

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

[7] G. Holton (1972), On trying to understand scientific genius, *American Scholar, 41,* 95-110.

[8] J.R. Josephson, B. Chandrasekaran, J.W. jr. Smith & M.C. Tanner (1986), Abduction by classification and assembly. In *PSA 1986*, vol. 1, Philosophy of Science Association, 458-470.

[9] S.M. Kosslyn (1980), *Image & Mind*, Cambridge, MA, Harvard University Press.

[10] S.M. Kosslyn (1983), *Ghosts in the Mind's Machine: Creating and Using Images in the Brain*, New York, W.W. Norton.

[11] S. M. Kosslyn & O. Koenig (1992), *Wet Mind. The New Cognitive Neuroscience*, New York, Free Press.

[12] L. Magnani (1992), Abductive reasoning: philosophical and educational perspectives in medicine. In D.A. Evans & V.L. Patel (Eds.), *Advanced Models of Cognition in Medical Training and Practice*, Berlin, Springer, 21-41.

[13] L. Magnani, S. Civita & G. Previde Massara (1994), Visual Cognition and Cognitive Modeling. In V. Cantoni (Ed.), *Human and Machine Vision: Analogies and Divergences*, New York, Plenum Press, 229-243.

[14] D. Marr & H.K. Nishihara (1978), Representation and recognition of the spatial organization of three-dimensional shapes. In *Proceedings of the Royal Society of London*, 269-294.

[15] A.I. Miller (1989), Imagery and intuition in creative scientific thinking: Albert Einstein's invention of the special theory of relativity. In D.B. Wallace & H.E. Gruber (Eds.) *Creative People at Work. Twelve Cognitive Case Studies*, Oxford, Oxford University Press, 171-187.

[16] N. J. Nersessian (1993), Opening the black box: Cognitive science and history of science. CSL Report 53, January 1993 (Cognitive Science Laboratory, Princeton University, Princeton, NJ).

[17] C.S. Peirce (1931-1958), *Collected Papers*, 8 vols, C. Harstone, P. Weiss & A. Burks (Eds.), Cambridge, MA, Harvard University Press.

[18] I. Peng & I.A. Reggia (1987), A probabilistic causal model for diagnostic problem solving I: integrating symbolic causal inference with numeric probabilistic inference, *IEEE Transactions on Systems, Man, and Cybernetics, 17*, 146-162.

[19] I. Peng & I.A. Reggia (1987), A probabilistic causal model for diagnostic problem solving II: diagnostic strategy. *IEEE Transactions on Systems, Man, and Cybernetics, 17*, 395-406.

[20] H.E. Pople (1973), On the mechanization of abductive logic. In *Proceedings of the International Joint Conference on Artificial Intelligence 8*, 147-152.

[21] Z.W. Pylyshyn (1981), The imagery debate: Analogue media versus tacit knowledge, *Psychological Review, 88*, 16-45.

[22] M. Ramoni, M. Stefanelli, L. Magnani & G. Barosi (1992), An epistemological framework for medical knowledge-based systems, *IEEE Transactions on Systems, Man, and Cybernetics, 22(6)*, 1361-1375.

[23] J.A. Reggia, S.N. Dana & Y.W. Pearl (1983), Expert systems based on set covering model, *International Journal on Man-Machine Studies, 19*, 443-460.

[24] P. Thagard (1988), *Computational Philosophy of Science*, Cambridge, MA, The MIT Press.

[25] P. Thagard (1992), *Conceptual Revolutions*, Princeton, NJ, Princeton University Press.

[26] P. Thagard, D. Gochfeld & S. Hardy (1992), Visual analogical mapping. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, Hillsdale, Erlbaum, 522-527.

[27] R.D. Tweney (1989), Fields of enterprise: On Michael Faraday's thought. In D.B. Wallace & H.E. Gruber (Eds.) *Creative People at Work. Twelve Cognitive Case Studies*, Oxford, Oxford University Press, 91-106.

[28] M. Tye (1991), *The Imagery Debate*, Cambridge, MA, The MIT Press.

HYPERACUITY: VERTICAL ASYMMETRY FOR SIZE DISCRIMINATION OF TWO-DIMENSIONAL IMAGES

Svetlana Chukova

Pavlov Institute of Physiology,

Emb.Makarov 6, 199034 St.Petersburg, Russia

The human eye is capable of resolving changes in the relative position of two visual features that are an order of magnitude smaller than the size of a foveal cone. To distinguish these highly precise spatial judgments from measures of visual acuity that assess resolution capability, Westheimer [1] coined the term hyperacuity. "Almost any type of target configuration can be used to measure hyperacuity" [2]: the exquisite sensitivities obtained in hyperacuity tasks are robust to changes in the spatial configuration of stimuli.

Can human observers judge size differences in two-dimensional images with the same exquisite precision? In this work, size discrimination thresholds for high-contrast, outline images were measured in the fovea for different arrangements of the stimuli along the central vertical meridian.

METHOD

The stimuli were black outline squares and rectangles on a bright background (60 cd/m^2). The reference stimuli were 27 squares with sides ranging in length from 8 to 60 arc min, in increments of 2 arc min. Five test rectangles were obtained by decreasing the height of the square by 1.25, 2.5, 3.75, 5.00 and 6.25% (thus, the test stimulus was always smaller than the reference). The outline width for both test and reference was 1/4 of the reference height.

The inspection field was limited by a circular aperture 7 deg in diameter. The reference and test stimuli were presented simultaneously along the central vertical meridian of the visual field (VF). The distance between the internal edges of the stimuli - 30 min arc - was constant for all the sizes. Viewing was central, with no explicit fixation target.

The order of the reference sizes was randomized. For each reference size, the five test-reference pairs were presented 100 times, 50 with the test above and 50 with the test below the re-