Hemispheric Asymmetry in Visual Perception Arises from Differential Encoding beyond the Sensory Level

Janet Hui-wen Hsiao (jhsiao@cs.ucsd.edu)
Reza Shahbazi (rshahbaz@ucsd.edu)
Garrison W. Cottrell (gary@cs.ucsd.edu)
Department of Computer Science & Engineering, University of California San Diego
9500 Gilman Dr. #0404, La Jolla, CA 92093 USA

Abstract

Hemispheric asymmetries in the perception of local and global features have been consistently reported: there is an advantage for responses to global features in the left visual field/right hemisphere and an advantage for responses to local features in the right visual field/left hemisphere. It has been proposed that this asymmetry originates from differential frequency bias in the two hemispheres (e.g., Ivry & Robertson, 1998). Nevertheless, there is little evidence supporting hemispheric specialization for particular frequency ranges (e.g., Fendrich & Gazzaniga, 1990) or differential frequency tuning in the neurons in the two hemispheres. Here we test the hypothesis that this hemispheric asymmetry in visual perception takes place at the encoding stage beyond the sensory level. We use two autoencoder networks with differential connectivity configurations as the way to develop differential encoding in the two hemispheres, to reflect the anatomical evidence that there is more interconnectivity among the neighboring cortical columns in the right hemisphere than the left hemisphere (e.g. Hutsler & Galuske, 2003). We show that this differential encoding mechanism has a better fit with human data than the model based on differential frequency bias, and thus is a more anatomically realistic and cognitively plausible model in accounting for the hemispheric asymmetry in visual perception.

Keywords: Hemispheric asymmetry, visual perception, Double Filtering by Frequency (DFF), autoencoder networks.

Introduction

The way we analyze and process the global and local forms of visual stimuli has been extensively examined. Navon (1977) proposed the "global precedence hypothesis", suggesting that the global form of a visual stimulus is unavoidably recognized before the local forms. This effect was later shown to depend on both the characteristics of the local and global forms and the hemispheric asymmetry in the perception of local and global features (Hoffman, 1980). Follow-up studies further confirm that there is a right visual field (RVF)/left hemisphere (LH) advantage for responses to local features and a left visual field (LVF)/right hemisphere (RH) advantage for responses to global features (Sergent, 1982; Ivry & Robertson, 1998). Nevertheless, studies examining grating detection did not support the existence of hemispheric specialization for particular frequency ranges (e.g. Di Lollo, 1981; Rijsdijk, Kroon, &Van der Wildt, 1980; Peterzell, Harvey, & Hardyck, 1989; Fendrich & Gazzaniga, 1990). It thus remains controversial about why this perceptual asymmetry exists (Peterzell, 1991; Martin, 1979).

Hierarchical letter Pattern Perception

Sergent (1982) used hierarchical letter patterns (Navon, 1977) to examine hemispheric differences in responses to global and local patterns. A hierarchical letter pattern contains two patterns: a global pattern and a local pattern. The global pattern (the large letter in Figure 1(a)) is composed of a number of local patterns (the small letters in Figure 1(a)). She referred to the two levels of the stimulus as having differential spatial frequency contents: low frequency for the global pattern and high frequency for the local pattern. In her experiment, she used four letters to compose the hierarchical letter patterns: "H" and "L" were designated as targets, and "T" and "F" as distracters. Given that each letter may appear as the local or the global pattern, there are in total 16 patterns, which can be put into six conditions according whether there is a target in the local or global patterns, as shown in Figure 1(a). Stimuli were presented to either the RVF/LH or the LVF/RH for 150 ms, and the participants' task was to judge whether they saw a target letter or not, regardless of its being in the global pattern or the local pattern.

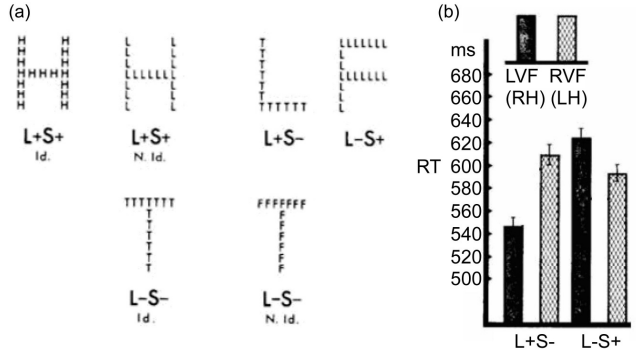

Figure 1: (a) Stimuli in Sergent's experiment (1982). "H/L" are targets, "T/F" are distracters. "L+" means the large letter is a target, and "S+" means the small letters are targets. "Id." means the local and the global patterns are identical. (b) The RT data for the L+S- and L-S+ stimuli in the LVF and RVF presentation conditions (Sergent, 1982).

The stimuli of greatest interest in Sergent's experiment (1982) were the conflict conditions when the target appeared in either the local pattern or the global pattern (i.e., the L+S- and L-S+ cases in Figure 1(a)), since they are the conditions in which interference may arise due to that the

target and the non-target letters are at different levels. The results showed that there was a significant interaction between the presented visual field (RVF or LVF) and the target level (global or local) in the response time data: the participants were faster at detecting the target at the global level when it was presented in the LVF/RH, and faster at detecting the target at the local level when it was presented in the RVF/LH (Figure 1(b)). She thus concluded that global precedence in form analysis (Navon, 1977) is a property of the RH but not the LH, and argued that this may result from a "greater capacity of the LH to deal with higher frequency". She also argued that "hemispheric differences as a function of spatial frequencies must result from processing taking place beyond the sensory level", since studies examining grating detection did not report a hemispheric difference in contrast sensibility or visible persistence (De Lollo, 1981; Rijsdijk et al., 1980. See also Peterzell et al., 1989; Fendrich & Gazzaniga, 1990); in other words, the two hemispheres may receive the same type of information beyond the sensory level (Sergent, 1982).

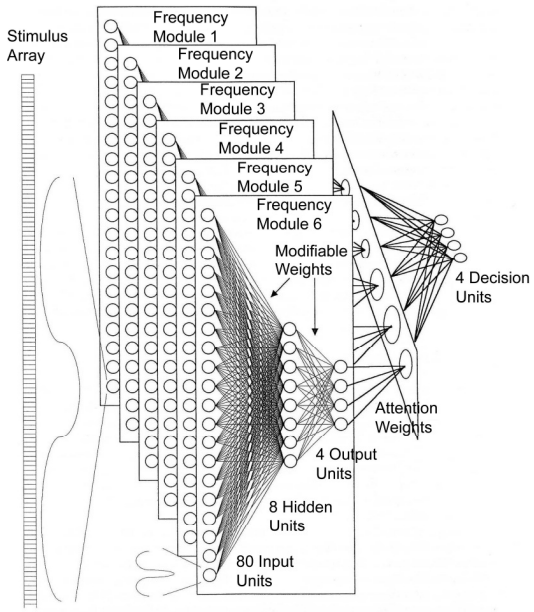

Figure 2: Ivry and Robertson's computational model of the DFF theory (taken from Ivry & Robertson, 1998). Module 6 has the lowest frequency. The four decision nodes correspond to four target patterns: whether target 1 or target 2 is present, and whether it is at the global or local level.

Modeling of the Double Filtering by Frequency (DFF) Theory

Ivry and Robertson (1998) proposed a Double Filtering by Frequency (DFF) theory to account for the hemispheric asymmetry in visual perception. The theory argues that information coming into the brain goes through two frequency filtering stages. The first stage involves attentional selection of task-relevant frequency information, and at the second stage the two hemispheres have asymmetric filtering processing: the LH amplifies high frequency information (i.e., a high-pass filter), whereas the RH amplifies low frequency information (i.e., a low-pass filter). They developed a computational model to account for this perceptual asymmetry effect (Figure 2). The model contains six different frequency modules; each module extracts information of a specific spatial frequency from the input and maps the extracted information to the output. The outputs from the modules then go through an "attention weight" layer as a filter. The filter first selects a frequency range that is expected to provide the most useful information for the task (although it is unclear how the range is decided). At the second stage, in the RH network, the filter amplifies the information from the low spatial frequency (LSF) modules within the frequency range, whereas in the LH network it amplifies the information from the high spatial frequency (HSF) modules, through giving different weights to different modules.

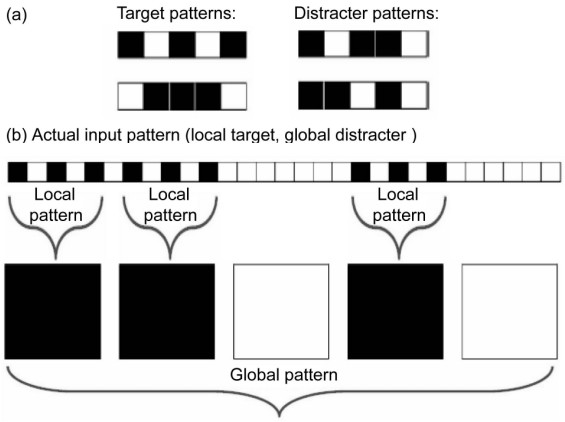

Figure 3: (a) In the input representation, there are two target patterns (10101 and 01110) and two distracter patterns (10110 and 11010). (b) Shown at the top is an actual input pattern formed by taking the second distracter pattern and replacing each black portion with the first target pattern. Thus this represents the first target pattern at the local level and the second distracter pattern at the global level (an L-S+ pattern). A 0 unit appears between each local pattern as a separator (cf. Ivry & Robertson, 1998).

The local input patterns in the simulation are created by turning on three bits in a five-bit vector. These are replicated three times, separated by all 0 bit patterns, to create the global pattern. The process is illustrated in Figure 3, where the global pattern is 11010. Each “1” in the global pattern is represented by the local pattern, 10101. Each “0” in the global pattern is represented by 00000.

The results showed that, consistent with the human data, the model exhibited the hemisphere-by-level interaction after 20 training epochs; nevertheless, this effect was not consistently obtained with more training epochs; the LH network was generally better for both local and global targets. In another simulation, they enlarged the stimuli by five so that each local pattern contained 25 units, with a group of five units corresponding to a unit in the original representation, and each global pattern thus became 125

units long. They showed that the hemisphere-by-level interaction was more pronounced in this simulation, and the interaction persisted across 100 training epochs (Figure 4). Unfortunately, this result is fragile: with further training, the LH network became better than the RH network at identifying both local and global targets, which is inconsistent with the human data. The model also exhibited a general advantage for a global target, which is consistent with the global precedence observation (Navon, 1977).

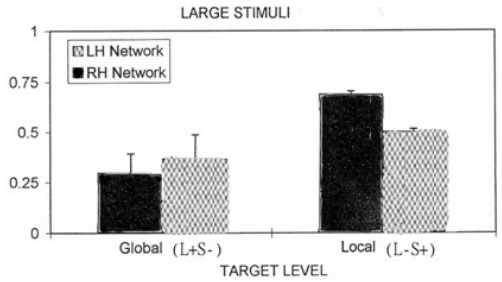

Figure 4: Results of Ivry and Robertson's computational model of the DFF theory (1998) with large stimuli after 100 epochs. Note that it is unclear whether the RH advantage in L+S- was significant. In addition, inconsistent with the human data, the LH network became better at identifying both local and global targets with further training.

Anatomical Differences between the Hemispheres

Although the DFF theory has been shown to be able to account for the observed perceptual asymmetry, there is little evidence suggesting differential frequency tuning in the neurons in the two hemispheres. Nor is there anatomical evidence supporting the existence of the different frequency modules in the brain as those proposed in Ivry & Robertson's computational model. In addition, given that it is unclear how the frequency range that is expected to provide the most useful information for the task is decided, Ivry and Robertson's model does not fully explain why there is little evidence suggesting the existence of hemispheric specialization for particular frequency ranges (e.g., Fendrich & Gazzaniga, 1990). Ivry & Robertson (1998) argued that it is because absolute instead of relative frequencies were used in these studies. As pointed out by Sergent (1982), given the lack of evidence of hemispheric specialization for particular frequency ranges, this hemispheric asymmetry must result from processing taking place beyond the sensory level. Contrary to the DFF theory, the two hemispheres may not differ in the way they extract visual information.

What is the process taking place beyond the sensory level that results in this perceptual asymmetry? Here we test the hypothesis that this process takes place at an encoding stage beyond the sensory level. We incorporate the anatomical evidence about hemispheric differences in cerebral cortical networks in a computational model that implements differential encoding in the two hemispheres. Galuske et al. (2000) examined the microcircuitry organization in human temporal cortex and showed hemispheric asymmetry in the posterior part of Brodmann area 22, which has been shown to be the area for language-relevant processing of auditory signals. They showed that, through neuronal track tracing, this area contains a modular network that links regularly spaced clusters of neurons; although the cluster size was similar in the two hemispheres, the spacing between the clusters was about 20 percent larger in the LH. This wider spacing implies that more functionally distinct columnar systems can be included per surface unit in the LH than in the RH. In addition, in the LH, each cortical column has fewer connections with neighboring columns compared with the RH (Seldon, 1981a; 1981b; 1982; Buxhoeveden et al., 2001). Taken together, these data suggest that there is more interconnectivity between the neighboring cortical columns of the RH than the LH, which may result in a more functionally overlapped representation in the RH compared with that in the LH (Hutsler & Galuske, 2003).

Although these anatomical data are from the auditory cortex because of the researchers' interests in language processing area, the phenomenon of differential frequency bias in the two hemispheres has been observed in both visual and auditory processing (Ivry & Robertson, 1998; Hutsler & Galuske, 2003). Thus, in the current study, we generalize the findings of hemispheric asymmetry in microcircuitry organization in the auditory processing area to visual processing, aiming to test the hypothesis that a difference in the connectivity configuration at the encoding stage of visual processing is sufficient to account for the observed hemispheric asymmetry in the perception of global and local features (the encoding scheme we propose here can be applied to auditory processing as well).

Modeling Method

To investigate the hypothesis that the observed perceptual asymmetry results from the difference in the connectivity configuration at the encoding stage, we use two autoencoder networks (Rumelhart, Hinton, & Williams, 1986) with different connectivity configurations as a way to learn an efficient encoding from the input data. An autoencoder network is a two-layer neural network trained to map the input pattern to an output pattern that is identical to the input pattern; after training, its hidden layer activation when an input pattern is presented is used as a compressed encoding of the input pattern. In our simulation, two autoencoder networks are created: the LH network has a pattern of connectivity resembling the sparse connections in the LH, whereas the RH network has more connections among the neighboring nodes to develop a more functionally overlapped representation. In the input representation, as in Ivry and Robertson's model (1998), we use two target patterns and two distracter patterns, each of which is five units long, to compose the hierarchical stimuli (Figure 3). Each stimulus is 29 units long, constructed by combining two patterns so that one pattern forms the local features and the other forms the global pattern of the stimulus, with a blank (0) unit between each local pattern. 16 stimuli are constructed from all possible combinations of the targets and distracters at either level (cf. Sergent, 1982).

Each autoencoder network has 13 hidden units; each hidden unit has 7 connections to the input layer. This parameter combination is the point where the overall error starts to converge and the two networks have a comparable performance level. A Gaussian probability density function (PDF) is used to determine the probability of having a connection between a hidden and the input units. For each hidden unit an identical Gaussian PDF is used, with the center of the Gaussian PDF located evenly within the input layer. A narrow Gaussian PDF is used for the RH autoencoder network ($\sigma = 1.8$), and a wide Gaussian PDF for the LH autoencoder network ($\sigma = 18$; see Figure 5)[1]; the variances are chosen as the two extreme cases of denseness/sparseness of the connections in order to examine the qualitative differences between them. The connections from the hidden layer to the output layer are completely symmetric to those from the input layer to the hidden layer.

After obtaining the efficient encoding of the input stimulus, we use a perceptron (i.e., a one-layer neural network) to classify the encoding according to whether there is a target (at either level) or not in the stimulus. Thus, the output layer of the perceptron has only one node; the node has value '1' when a target is present and '0' otherwise. The error was measured as the difference between the output of the perceptron and the desired output (0 or 1).

For both the autoencoder networks and the perceptron, the training algorithm is gradient descent with back-propagation (Rumelhart et al., 1986). The initial learning rate is 0.05 for the autoencoder networks and 0.001 for the perceptron, and is flexible during training: if the error decreases in the current epoch, the learning rate for the next epoch is the old learning rate multiplied by 1.01; if the error increases, the new learning rate is the old learning rate divided by two. The training stops when the number of training epoch reaches 1,000 for the autoencoder networks and 1,500 for the perceptron; the numbers of epochs are chosen so that the error will have converged before the training stops.

Results

Figure 6(a) shows the mean error over 500 trials for each network (typically a correspondence is made between reaction time and error – more uncertainty leads to longer reaction time)[2]. The results showed a significant interaction between network (RH vs. LH) and target level (global vs. local) ($F(1, 973) = 53.140$, $p << 0.001$), which is consistent with human reaction time data (Sergent, 1982): an advantage of responses to a target at the global level in the RH network compared with the LH network ($F(1, 973) = 8.170$, $p < 0.01$), and an advantage of responses to a target at the local level in the LH network compared with the RH network ($F(1, 979) = 7.652$, $p < 0.01$). In addition, similar to Ivry and Robertson's model (with large stimuli), the results also showed an overall advantage for responses to a global-level target ($F(1, 973) = 53.484$, $p << 0.001$), which is consistent with the global precedence of visual form analysis (Navon, 1977). There was no main effect of network architecture ($F < 0.1$, *n.s.*), which is also consistent with the human data (Sergent, 1982; Figure 6(b)).

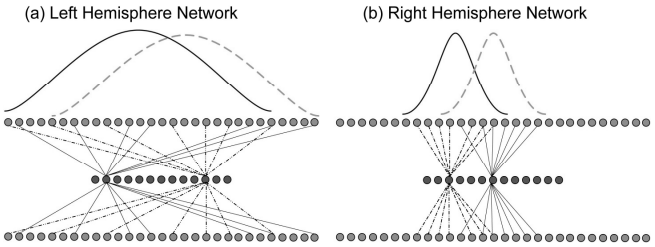

Figure 5: (a) Probability density function of the connections in the LH network ($\sigma = 18$) and an example LH network. (b) Probability density function of the connections in the RH network ($\sigma = 1.8$) and an example RH network.

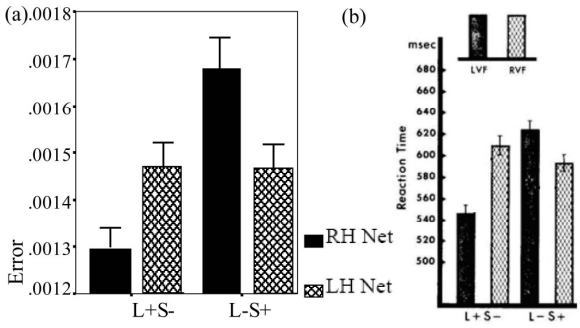

Figure 6: (a) Modeling results. The error bars show standard errors. L+S- means the target is at the global level, whereas L-S+ means the target is at the local level (Figure 1). (b) Sergent's experiment results (1982), duplicated here for comparison reasons.

The combination of 13 hidden nodes and 7 connections from each hidden node was chosen in our simulation because it is the point where the error of the perceptron started to converge. We also explored how the network performance changes with different parameters. We found that, while keeping the number of connections from each hidden node constant, with 14 hidden nodes the two networks had comparable performance level ($F(1, 961) = 0.005$, *n.s.*), and the interaction between target level and network was significant ($F(1, 961) = 22.100$, $p << 0.001$) – a result very similar to the networks with 13 hidden nodes. With more than 14 hidden nodes (i.e., 15) or fewer than 13 (i.e., 11 and 12), the interaction between target level and network remained significant, but the RH networks outperformed the LH networks when there were more than

[1] Note that this Gaussian PDF is used to create *connections between layers* and thus is different from the Gaussian receptive field functions used in some models of hemispheric asymmetry (e.g. Ivry & Robertson, 1998; Monaghan & Shillcock, 2004). Also, the distribution widths are opposite of ours (e.g., wide right).

[2] Since the connectivity configuration for each trial was randomly generated, in some rare cases the network did not learn well and ended up with extremely large error. We use the mean plus/minus two standard deviations as the upper/lower bound to remove the outliers, which is about 2 to 3% of the simulation trials.

14 hidden nodes, whereas the LH networks outperformed the RH when there were fewer than 13 hidden nodes. In contrast, while keeping the number of hidden nodes constant, the interaction between target level and network also remained significant, but the RH networks outperformed the LH networks with more than 7 connections (i.e., 8 and 9), whereas the LH networks outperformed the RH networks when there are fewer than 7 connections (i.e., 5 and 6). This result shows that the interaction between target level and network architecture is robust against parameter changes; changes in the number of hidden nodes and the number of connections from each hidden node only alter the difference in performance level between the two networks. The model fits the human data the best when the two networks have a similar performance level; this is also consistent with the human data that there is no main effect of hemisphere in performance.

We also explored how the performance changed with the variance on the connection distribution; this allowed us to determine the proper parameters for the LH and RH models. Figure 7 shows how the network behavior changes with the density of the connections in the model with 13 hidden nodes and 7 connections from each hidden node. The far left (1.8) is the RH network and the far right (18) is the LH network. It is clear that the interaction would survive less extreme parameter variation.

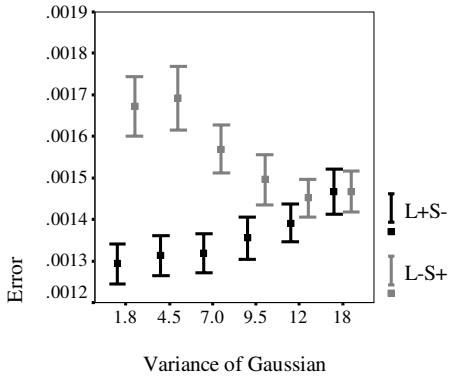

Figure 7: The network's performance with different connection density: the larger the variance of the Gaussian distribution, the sparser the connection configuration is.

Compared with Ivry and Robertson's model, the current model has a better fit with the human data (Sergent, 1982). In their results, a better match with the human data was obtained when they used large stimuli. However, even with the large stimuli, it is unclear whether the hemispheric difference was significant in each level condition as the human data (they only reported a significant interaction between network and level but not for each level condition separately; see Figure 4); also, the LH network eventually became better at identifying targets at both levels, which is inconsistent with the human data. In contrast, our results better match the human data (Figure 6): a significant hemispheric difference in both level conditions, and a comparable performance between the two hemispheres.

Conclusion & Discussion

Here we test the hypothesis that the hemispheric asymmetry in the perception of global and local features originates from differential encoding beyond the sensory level, instead of differential frequency bias as proposed by the DFF theory (Ivry & Robertson, 1998). We first argue that the lack of evidence supporting hemispheric specialization for particular frequency ranges (e.g., Di Lollo, 1981; Peterzell et al., 1989; Fendrich & Gazzaniga, 1990) suggests that this asymmetry takes place beyond the sensory level (Sergent, 1982), and the two hemispheres do not differ in information extraction. We then argue that the difference takes place at the encoding stage. We incorporate evidence about the differences in the anatomical and functional microstructure of the two hemispheres in a computational model that uses autoencoder networks to develop efficient encoding of the stimulus (Rumelhart et al., 1986): the columnar structure in the RH has more interconnectivity among neighboring columns compared with that in the LH, and thus may be more functionally overlapped than that in the LH (Hutsler & Galuske, 2003). We thus use two autoencoder networks with different connectivity configurations to simulate this differential encoding: the RH network has a narrower connectivity distribution to allow more connectivity among neighboring nodes compared with the LH network. We then use a perceptron network to examine how good the two encoding systems are in terms of detecting local and global level targets. The results match the human data very well; it shows a significant hemisphere-by-level interaction: a RH advantage for responses to a global level target, and a LH advantage for responses to a local level target (Sergent, 1982); and an overall advantage in responses to a global level target (Navon, 1977).

In the comparison with Ivry and Robertson's model (1998) based on the DFF theory, we show that our model has a better fit with the human data (Sergent, 1982). Their model enforces a discrete separation of frequency information into modules and the hemispheric differences take place through a manipulation of the combination of the different frequency modules; it is thus unclear how these frequency ranges are selected and why they are combined in a certain way, and how the model is able to account for the lack of evidence supporting hemispheric specialization for particular frequency ranges (e.g., Fendrich & Gazzaniga, 1990). In contrast, through hypothesizing that hemispheric differences take place at the encoding stage, and using the Gaussian distribution to simulate differential connectivity configuration at the encoding stage, our model develops naturally the hemispheric difference in the frequency content in the encoding, without assuming two stages of frequency filtering (Ivry & Robertson, 1998). In addition, there is little anatomical evidence suggesting differential frequency tuning in the neurons in the two hemispheres, or different frequency modules in the brain similar to those proposed in their model. In short, compared with their model, our model is more anatomically realistic, more cognitively plausible, and has a better fit to the human data.

In the current simulation we use the same stimuli as those used in Ivry and Robertson's model (1998) for comparison reasons. Although the stimuli are simpler than those used in the human experiments (Sergent, 1982), the current simulation presents an important clue in the debate of this perceptual asymmetry that has puzzled researchers for at least 25 years: this difference may be due to differential encoding in the two hemispheres beyond the sensory level; computational modeling makes it possible to demonstrate this differential encoding effect. It also predicts that a hemispheric difference in cortical column structure similar to that in the auditory cortex may also exist in the high-level visual areas such as inferior temporal cortex.

We are currently pursuing the incorporation of more anatomical data into the model, such as using 2D Gabor filters to simulate responses of complex cells in the early visual system (Daugman, 1985), and also using the proposed autoencoder networks as the way to develop efficient encoding in the two hemispheres in modeling more complicated (e.g. the hierarchical letter patterns) and real world visual stimuli (such as faces; cf., the Principal Component Analysis step in many visual perception models), in order to examine the cognitive plausibility of this differential encoding mechanism in accounting for various hemispheric asymmetry phenomena in perception, such as the left side bias effect in face perception (e.g., Gilbert & Bakan, 1973) and the RVF advantage in visual word recognition (e.g., Bryden & Rainey, 1963). This differential encoding account may also be applied to hemispheric asymmetry in other cognitive domains. For example, Monaghan and Shillcock (2004) use different receptive field sizes in the two hemispheres, similar to the different frequency modules in the DFF model, in a computational model of visual attention aiming to account for the unilateral visual neglect phenomenon. Their model also suffers from the lack of evidence for differential frequency biases in the hemispheres at the sensory level; whether the differential encoding mechanism can also account for this asymmetry in visual attention is currently under examination.

Acknowledgments

This research was supported by NIH grant MH 57075 and NSF grant #SBE-0542013 to the Temporal Dynamics of Learning Center (GWC, PI) and a McDonnell Foundation grant to the Perceptual Expertise Network (I. Gauthier, PI).

References

Bryden, M.P., & Rainey, C.A. (1963). Left-right differences in tachistoscopic recognition. *Journal of Experimental Psychology*, *66*, 568–571.

Buxhoeveden, D. P., Switala, A. E., Litaker, M., Roy, E., & Casanova, M. F. (2001). Lateralization of minicolumns in human planum temporal is absent in nonhuman primate cortex. *Brain Behav. Evol*., *57*, 349-358.

Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A*, *2*, 1160-1169.

Di Lollo, V. (1981). Hemispheric symmetry in visible persistence. *Perception & Psychophysics*, *11*, 139-142.

Fendrich, R., & Gazzaniga, M. (1990). Hemispheric processing of spatial frequencies in two commissurotomy patients. *Neuropsychologia, 28,* 657–663.

Galuske, R. A., Schlote, W., Bratzke, H., & Singer, W. (2000). Interhemispheric asymmetries of the modular structure in human temporal cortex. *Science*, *289*, 1946-1949.

Gilbert, C. & Bakan, P. (1973). Visual asymmetry in perception of faces. *Neuropsychologia*, *11*, 355-362.

Hoffman, J. E. (1980). Interaction between global and local levels of a form. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 222-234.

Hutsler, J. & Galuske, R. A. W. (2003). Hemispheric asymmetries in cerebral cortical networks. *Trends in Neurosciences*, *26*, 429-435.

Ivry, R. & Robertson, L. C. (1998). *The Two Sides of Perception*. Cambridge: MIT Press.

Martin, M. (1979). Hemispheric specialization for local and global processing. *Neuropsychologia*, *17*, 33-40.

Monaghan, P. & Shillcock, R. C. (2004). Hemispheric asymmetries in cognitive modeling: connectionist modeling of unilateral visual neglect. *Psychological Review*, *111*, 283-308.

Navon (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353-383.

Peterzell, D. H. (1991). On the nonrelationship between spatial frequency and cerebral hemispheric competence, *Brain & Cognition, 15,* 62-68.

Peterzell, D. H., Harvey, L. O., Jr. & Hardyck, C.D. (1989). Spatial Frequencies and the cerebral hemispheres: Contrast sensitivity, visible persistence, and letter classification. *Perception & Psychophysics, 46, 433-455.*

Rijsdijk, J. P., Kroon, J. N., &Van der Wildt, G. J. (1980). Contrast sensitivity as a function of position on retina. *Vision Research*, *20*, 235-241.

Rumelhart, D. E., Hinton, G. E, & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*, 533-536.

Seldon, H. L. (1981a). Structure of human auditory cortex. I. Cytoarchitectonics and dendritic distributions. *Brain Research*, *229*, 277–294.

Seldon, H. L. (1981b). Structure of human auditory cortex. II. Axon distributions and morphological correlates of speech perception. *Brain Research*, *229*, 295–310.

Seldon, H. L. (1982). Structure of human auditory cortex. III. Statistical analysis of dendritic trees. *Brain Research*, *249*, 211–221.

Sergent, J. (1982). The Cerebral Balance of Power: Confrontation or Cooperation? *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 253-272.