PERGAMON

Vision Research 42 (2002) 2547–2554

Vision Research

www.elsevier.com/locate/visres

The role of complex cells in object recognition

Ladan Shams [a,*], Christoph von der Malsburg [b,c]

[a] *Division of Biology, California Institute of Technology, MC 139-74 Pasadena, CA 91125, USA*
[b] *Institut für Neuroinformatik, Ruhr-Universität Bochum, D-44780 Bochum, Germany*
[c] *Computer Science and Neuroscience Departments, University of Southern California, HNB 228, Los Angeles, CA 90089-2520, USA*

Received 27 August 2001; received in revised form 24 June 2002

Abstract

Primate's primary visual cortex (V1) is dominated by complex cells. This choice of nature seems puzzling, as complex cells are insensitive to spatial phase—information which is generally believed to be essential for perceptual characterization and recognition of images. Modeling complex cells as Gabor wavelet magnitudes, we have mathematically and empirically examined the information content of their responses. Our results show that in spite of phase insensitivity of individual complex cell responses, population responses contain sufficient information to capture the perceptual essence of images. A complex cell type representation seems to be not only sufficiently discriminating for object identification, but also—due to its inherent ambiguities—robust to changes in background, lighting, and small deformations.

Keywords: Complex cell; Phase; Recognition; Gabor; Object representation

1. Introduction

Fourier components of an image can be expressed in terms of magnitude and phase. It is commonly believed that the phase information dominates perception of visual scenes (Oppenheim & Lim, 1981; Palmer, 1999; Piotrowski & Campbell, 1982; Shapley, Caelli, Grossberg, Morgan, & Rentschler, 1990). This notion is based on demonstrations such as those illustrated in Fig. 1. Images obtained by inverse Fourier transform (IFT) of correct magnitudes (or power spectra) but with arbitrary (constant or random) phase spectra are not recognizable as the scene in the original image; in fact, they look garbled (e.g., Fig. 1c). This demonstration seems to attest the importance of phase information in the perception of scenes. Furthermore, hybrid images obtained by IFT of the magnitudes of one image combined with phases of another image are invariably perceived as the image where the phase spectra came from (Fig. 1e,f). This observation has led to the belief that phase information is more important than magnitude information in perceiving objects and scenes.

On the other hand, primate's V1 is dominated by complex cells (Dow, 1974; Hubel & Wiesel, 1968; Poggio, 1972; Schiller, Finlay, & Volman, 1976), i.e., cells whose responses are characterized by selectivity to orientation and frequency while lacking sensitivity to the *spatial phase* of gratings (De Valois, Albrecht, & Thorell, 1982; Movshon, Thompson, & Tolhurst, 1978; Pollen & Ronner, 1982; Skottun et al., 1991). Most area V4 cells were also found to be insensitive to spatial phase and local stimulus position and thus to be more similar to V1's complex cells rather than simple cells (Gallant, Braun, & Essen, 1993). These facts are puzzling since the neglect of important information on phase and local position in a vast class of neurons should create immense image signal ambiguities and thus disrupt perceptual processes. Why should phase information, which seemingly plays a crucial role in our visual perception, be ignored by a sweeping majority of cells at the most fundamental level of visual processing? We have investigated the functional implications of a scene representation based on complex cells, and the results of our study offer a resolution to this puzzle.

The goal of our study is to examine how "good" a scene representation based on complex cell signals can be for the task of scene recognition. In other words, we would like to examine how much and what kind of

---

[*] Correponding author. Tel.: +1-626-395-2362; fax: +1-626-844-4514.
E-mail address: ladan@caltech.edu (L. Shams).

0042-6989/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0042-6989(02)00202-X

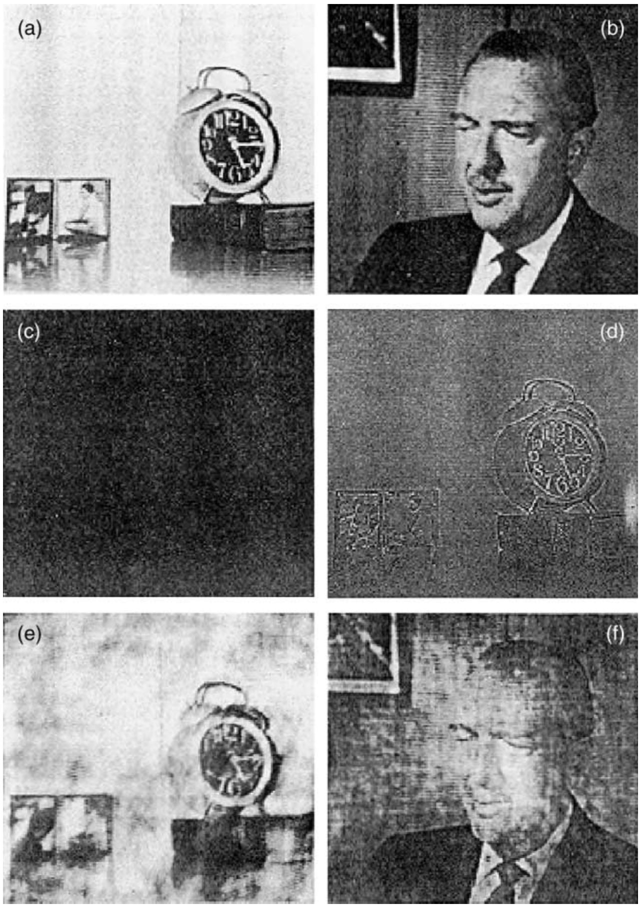

Fig. 1. IFT of magnitude and phase. (a) and (b) Two arbitrary images and (c) the image obtained from IFT of magnitudes of image in (a) combined with constant (zero) phase. A similar result, that is a meaningless image, would be obtained if random phases were used instead of constant phase. (d) The result of the IFT of phases of image (a) combined with constant (unit) magnitude. These observations appear to suggest that unlike phase, the magnitude information is not sufficiently rich to capture the perceptual essence of images. (e) The result of IFT of the phases of (a) and magnitudes of (b). (f) The result of IFT of the phases of (b) and magnitudes of (a). These results seem to suggest that the phase information is dominant in determining the perceptual essence of the scenes. (Images adopted from Oppenheim & Lim, 1981.)

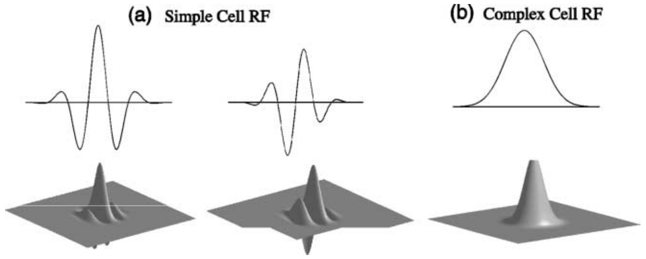

Fig. 2. Schematics of simple and complex cell RF profiles and their corresponding models. (a) The top drawings depict the response profiles (space domain) of two simple cells to a flashing bar of a given orientation and width in different spatial positions. Below, the real and imaginary parts of the Gabor wavelet profile of a given orientation and frequency is illustrated simulating the simple cell RFs. The real (or even-symmetric) and imaginary (or odd-symmetric) parts, different only in phase, are shown on left and right, respectively. (b) The top drawing depicts a complex cell response profile to the same type of stimuli. Below, the profile of the magnitude of the Gabor wavelets shown in (a), used as a model of the complex RF, is shown. While complex cells share the orientation and frequency tuning characteristics of simple cells, their RF is more spatially extended and smooth than those of simple cells. These traits are accurately modeled by Gabor magnitudes.

information is retained or lost at this stage of visual processing. To this end, we model complex cell responses with magnitude of Gabor wavelets, as described below:

The receptive fields (RFs) of V1 simple cells can be interpreted as Gabor wavelet functions (Daugman, 1985; Jones & Palmer, 1987; Marcelja, 1980) (Fig. 2a). A Gabor wavelet (Grossmann & Morlet, 1985) is a sinusoidal wave enveloped by a Gaussian. Gabor components of an image—obtained by convolution of the image with Gabor kernels—can be expressed in terms of magnitude and phase. Gabor magnitudes are tuned to a specific orientation and frequency and lack spatial phase selectivity. Therefore, the response profile of complex cells can be interpreted as Gabor magnitudes (Daugman, 1993; Daugman & Downing, 1995) (Fig. 2b). A qualitative comparison of our Gabor magnitude profiles and the experimental data on complex cell RF profiles (Ohzawa, DeAngelis, & Freeman, 1997) clearly confirms that Gabor magnitudes provide a good approximation to complex cell RFs. Our model of complex cell RF is a quite standard one, based on the numerous previous models (Adelson & Bergen, 1985; Daugman, 1993; Field, 1987; Morrone & Burr, 1988; Pollen, Gaska, & Jacobson, 1988; Spitzer & Hochstein, 1988) which have used sum of squared [1] outputs of quadrature-phase local (simple-cell-like) filter pairs (Pollen & Ronner, 1981).

To examine the characteristics of a complex cell representation, we need to find how ambiguous such a representation is. The more ambiguous a code is the larger the number of images which could give rise to it, and thus, the less discriminative it will be. The *kinds* of ambiguities are also very informative as they determine the type of invariances that a representation provides. An optimal representation for the recognition system would be one that retains enough information to discriminate between different objects, while containing ambiguities which would provide invariance with respect to changes that the projections of a given object may undergo (e.g., changes due to lighting, background, etc.).

From a complex cell's response it is impossible to determine whether the stimulus is an edge or a line, whether it is a black on white or white on black line,

---

[1] Or similar other non-linearities.

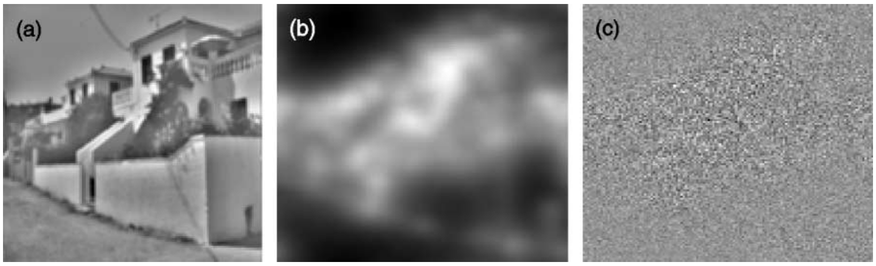

Fig. 3. IGT of magnitude combined with arbitrary phase. (a) An arbitrary image, (b) and (c) the result of IGT of the magnitudes of image (a) combined with zero phases, and random phases, respectively. These images are not perceptually meaningful, much less do they bear any resemblance to the scene in image (a).

whether it is a black-white or a white-black edge, and where the line or edge is positioned within the RF (a bright bar away from the RF center would cause the same response as one dimmer but closer to the center). These ambiguities exist for every single complex cell response, i.e., at every retinotopic coordinate, and for every orientation and frequency tuning. Thus, the ambiguity of this representation for the entire scene can potentially be combinatorial!

2. Reconstructing images from complex cell responses

In order to examine the ambiguity of the Gabor magnitude representation for an entire scene one can try reconstructing images from that representation. Reconstruction of images based merely on magnitude information, however, is not a trivial task. The inverse Fourier images shown in Fig. 1 have been interpreted as reconstructions from Fourier magnitude. A similar method can be applied to Gabor magnitudes. Fig. 3b and c display images that are derived from inverse Gabor transforms (IGTs) of the correct magnitudes of an image combined with arbitrary set of phases. As can be seen, the results are similar to those of inverse Fourier transforms; the images look garbled. It should be noted, however, that neither these images nor those in Fig. 1 are true reconstructions [2] of the magnitude information, as they are computed based not merely on magnitude information but rather, the magnitude information combined with wrong phase information, and the magnitudes of the forward transform of these images are not equal to the original magnitudes. The fundamental problem with this method is that the missing set of phases are not derived, but rather set a priori to arbitrary values. In order to bypass this problem and obtain valid reconstructions of Gabor magnitude representations we have developed a method for reconstructing images exclusively from magnitude information.

In a nut-shell, the algorithm works as follows. Starting with an arbitrary seed image as a first rough estimate of the reconstruction, we iteratively update the estimate such that its Gabor magnitudes become more and more similar to the target's Gabor magnitudes. This process converges to an estimated image which has the correct magnitudes, and can hence be considered a valid reconstruction. Depending on the seed image the algorithm can lead to different reconstructions. The algorithm is described in detail below.

2.1. Gabor wavelet transform

Given an image with gray values $I(\vec{\mu})$ defined on a two-dimensional lattice of pixel positions $\{\vec{\phi}\}$, the Gabor transform (or more accurately, Morlet transform) is defined as

$$J_{\vec{k}\vec{\mu}} = \sum_{\vec{\mu}'} I(\vec{\mu}')\psi_{\vec{k}}(\vec{\mu} - \vec{\mu}'), \quad \vec{\mu}, \vec{\mu}' \in \{\vec{\phi}\}$$

$$\text{with} \quad \psi_{\vec{k}}(\vec{\mu} - \vec{\mu}') = \frac{\vec{k}^2}{\sigma^2} \mathrm{e}^{-\frac{\vec{k}^2(\vec{\mu}-\vec{\mu}')^2}{2\sigma^2}} \mathrm{e}^{i\vec{k}(\vec{\mu}-\vec{\mu}')}. \quad (1)$$

Vectors $\vec{k}$ and $\vec{\mu}$ specify the orientation and frequency tuning, and the spatial coordinates of the wavelet, respectively. As the DC component of the Gabor transform is negligible, removing it by an additional term, as proposed originally by Morlet, does not affect the results of reconstruction. We here assume a sampling of the image domain at pixel density and of the frequency domain at a set F of frequency levels and, within a frequency level, at a fixed set O of orientations, $\vec{k} \in F \times O$. Set F consists of seven spatial frequency levels (with highest angular frequency equal to $3\pi/4$) at octave intervals, and O consists of 12 orientations equally spaced between 0 and π. The value of σ is set to π, leading to RF

---

[2] This statement applies to all images that are derived by applying the inverse Fourier (or Gabor) transform to the correct amplitude information combined with arbitrary phase information or vice versa, as in (Morgan, Ross, & Hayes, 1991; Oppenheim & Lim, 1981; Piotrowski & Campbell, 1982; Shapley et al., 1990). The results of these inverse transforms cannot be considered valid reconstructions (Hayes, 1982) of magnitudes (or phases) as the forward transform of these "reconstructions" will not produce the original magnitudes (or phases).

standard deviation sizes of 1.3 and 10.6 pixels for the smallest and largest RFs, respectively. Image resolution is 64 × 64 pixels. Gabor components can be expressed in terms of magnitude and phase, $J_{\vec{k}\vec{\mu}} = |J_{\vec{k}\vec{\mu}}|e^{i\phi_{\vec{k}\vec{\mu}}}$. We refer to the magnitudes $|J_{\vec{k}\vec{\mu}}|$ as Gabor magnitudes.

2.2. Reconstruction algorithm

Let $I(\vec{\mu})$ and $|J_{\vec{k}\vec{\mu}}|$ be an image and its Gabor magnitudes, respectively. Starting with an arbitrary seed image, we iteratively compute a sequence of images $\widehat{I}(\vec{\mu})$ whose Gabor magnitudes $|\widehat{J}_{\vec{k}\vec{\mu}}|$ approximate those of I with higher and higher accuracy. To this end we minimize the error

$$E = \sum_{k\vec{\mu}} (|J_{\vec{k}\vec{\mu}}|^2 - |\widehat{J}_{\vec{k}\vec{\mu}}|^2)^2 \tag{2}$$

with the help of gradient descent, $\partial\widehat{I}_{\vec{\mu}}/\partial t = -\partial E/\widehat{I}_{\vec{\mu}}$. Taking this gradient of Eq. (2) and mathematically manipulating the result leads to the update rule

$$\Delta\widehat{I}_{\vec{\mu}} = \xi \sum_{k} \left(\sum_{\vec{\mu}'} D_{\vec{k}\vec{\mu}'} \widehat{J}_{\vec{k}\vec{\mu}'} \psi_{\vec{k}}(\vec{\mu} - \vec{\mu}') \right)^{R} \tag{3}$$

with the abbreviation $D_{\vec{k}\vec{\mu}'} = |J_{\vec{k}\vec{\mu}'}|^2 - |\widehat{J}_{\vec{k}\vec{\mu}'}|^2$, R denoting the real part and ξ a parameter to scale the update speed. In each iteration, all pixels are updated according to Eq. (3). As mentioned before, the gradient descent is initialized with a seed image, for which we have experimented with uniform and Gaussian noise as well as photographs of natural scenes, finding the latter to lead to faster convergence. Fig. 4 illustrates a flow-chart diagram of the reconstruction algorithm. For a quantitative evaluation of the quality of the reconstruction we use the median over $\vec{\mu}$ and $\vec{k}$ of relative errors $e_{\vec{k}\vec{\mu}} = \frac{(\mathrm{abs}(|J_{\vec{k}\vec{\mu}}| - |\widehat{J}_{\vec{k}\vec{\mu}}|))}{|J_{\vec{k}\vec{\mu}}|}$ (avoiding noise due to small denominators by excluding small magnitude values from this computation). The iteration is stopped when this median error falls below a threshold (0.05). We set the step size ξ as large as compatible with a reduction of error (2). The necessary number of iterations is typically several hundred.

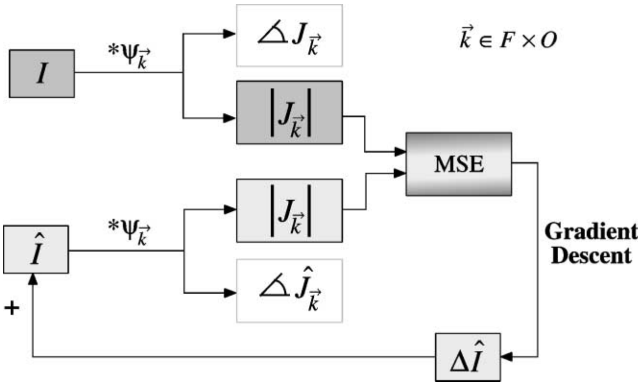

Fig. 4. The iterative algorithm used to reconstruct target image I from its Gabor magnitudes $|J_{\vec{k}}|$. An arbitrary seed image is used in the first iteration as $\widehat{I}$ to get the iteration started. The trial image $\widehat{I}$ is transformed with the same Gabor wavelets $\psi_{\vec{k}}$ as the target image. Transformations result in a magnitude and a phase component at each image coordinate for each wavelet. The algorithm only uses the magnitudes for reconstruction. At each iteration, the mean squared error between the squared magnitudes of the target and trial image transforms, $|J_{\vec{k}}|$ and $|\widehat{J}_{\vec{k}}|$, respectively, are computed, and using the gradient descent method an update $\Delta\widehat{I}_{\vec{\mu}}$ to the trial image is calculated to decrease the Gabor magnitude error. $\widehat{I}$ will be the reconstruction of I when the iteration stops, i.e., when the $|\widehat{J}_{\vec{k}}|$ have converged to $|J_{\vec{k}}|$. It should be noted that the image reconstruction method is used merely as a tool for us to examine the information content of the complex cells responses, and by no means should it be taken as a suggestion that brain engages in image reconstruction of the complex cell responses. Therefore, biological plausibility of the reconstruction method is irrelevant.

3. Results

Reconstructions of three images are shown in Fig. 5. The target images are shown on the left column. For each target image, two reconstructions are displayed in the middle and right columns. Considering the observations of Figs. 1 and 3 and the existing dogma about the importance of phase information, the coherence and recognizability of the reconstructions is quite surprising. We have computed a gray-level pixel error for each reconstruction (see figure caption). While the Gabor magnitude error for all reconstructions is the same (negligible), the pixel (and phase) error varies from one reconstruction to another for a given target, and in some cases this error is quite large. These variations reflect the ambiguity of the representation. Reconstructions differ considerably from each other, varying in contrast polarity and/or shading of local regions or of the entire image. A closer look also reveals small shifts in the location of some structures (e.g., lines and edges). Nevertheless, to our eye the essence of the original is rendered with ample clarity to recognize the objects.[3]

Our reconstructions show that the ambiguities inherent in the responses of individual complex cells are

---

[3] When we tightened the permissible error margin, median $e_{\vec{k}\vec{\mu}}$, to a very small number—in effect not allowing inaccuracy in more than a couple of pixels—only two types of reconstructions were left: fairly accurate replicas of the original images or of their negatives. This shows that there are only two attractors in this optimization problem, or in other words, there are two solutions to the problem of reconstruction. Intriguingly, however, the results obtained with a less strict error threshold (shown and discussed above) show that these two attractors are surrounded by images that are all perceptually meaningful. Given the many sources of data variation from image to image, and general noisiness of biological neurons' firing behavior, analysis of complex cell representation cannot be restricted to the exact mathematical solutions (the two attractors) and must consider the small neighborhood surrounding these attractors as solutions.

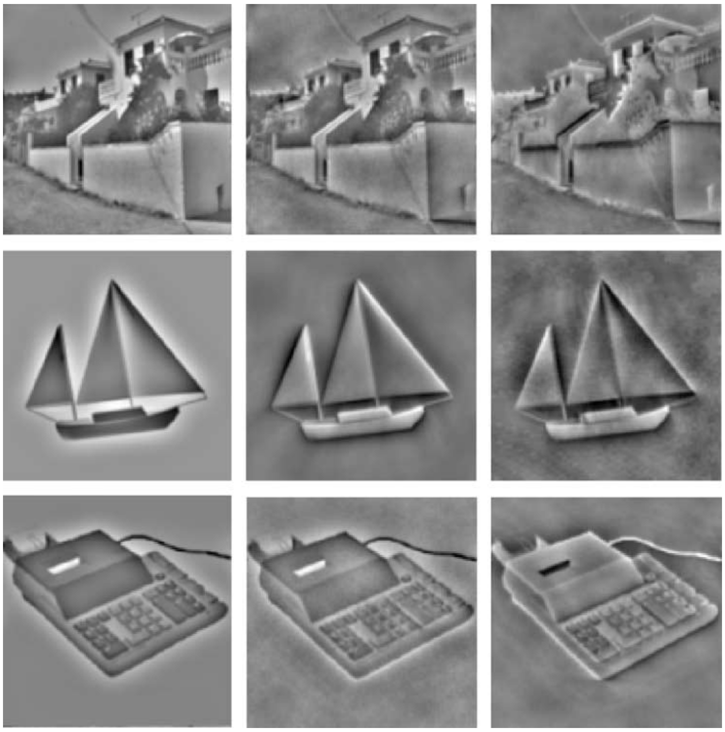

Fig. 5. Reconstruction from Gabor magnitudes. The images in the left-most column are the target images whose Gabor magnitudes have been used for reconstruction. These images are filtered in order to be limited to the same bandwidth visible to the Gabor wavelets. For each target image, we have displayed two different reconstructions in the middle and right columns. The pixel errors (defined as a counterpart of $e_{\vec{k}\vec{\mu}}$ but for the pixel values) vary tremendously between reconstructions. The following are the median pixel errors for the reconstructions of each target image, with the first and second numbers corresponding to the left and right reconstructions, respectively. These numbers signify the size of the error for individual pixels compared to the actual pixel value, e.g., an error of 0.5 means that on average the error in each pixel is half (or 50%) of the actual pixel value. Building: 0.16 and 1.52. Boat: 1.71 and 1.46. Calculator: 0.21 and 1.67.

coupled by their overlap both in the spatial and the frequency domain. The pixel values of a reconstruction have to reproduce all Gabor magnitudes simultaneously. For example, in the reconstruction of two regions corresponding to two neighboring points of an edge, if one was inverted in contrast, the superposition of the two would lead to a very small signal, all but wiping out the magnitudes for both. Thus, the reconstruction would have to make a choice and give the edge one of the two possible polarities on which both reconstructions would have to agree. It is to be expected, however, that the strength of this coupling depends on the actual image content—two image regions separated by blank space or an empty frequency band would have no reason to agree in contrast polarity, for instance. It is also to be expected that there are strongly coupled image domains, in which a consistent reconstruction will arise, different domains making independent contrast polarity or exact position choices. Domains may be limited both in their spatial and in their frequency extent.

4. Discussion

We examined the information content of an image's complex cell representation for the purpose of object recognition by trying to reconstruct the original image from the complex cell responses. We would like to emphasize that we are not suggesting that image reconstruction is a mechanism used by the brain. We have used reconstruction merely as a tool to analyze the information content of the representation under question: to examine how much and what kind of information about the scenes may be lost, whether the loss of information would amount to disruption of discriminability, and what kind of invariances may be provided by the ambiguities inherent in the representation.

4.1. Richness of complex cell representation

Our reconstruction results clearly show that, despite the absence of phase information, complex cell responses are sufficiently rich to represent objects and scenes for identification. Although intuitively it may be expected that infinitely many meaningless images (such as those shown in Fig. 3)—one image for each set of random phases—could give rise to a given set of Gabor magnitudes, our results prove that this intuition is simply wrong. The fact that all reconstructions of a given image are meaningful and readily recognizable as the original image clearly demonstrates that the ambiguity of this representation is limited to the set of images which share their perceptual "meaning". This implies that only very special phase relationships can give rise to a given set of magnitudes, thus, in the aggregate, the ambiguity of a population of complex cells is much less than the sum of ambiguities of the individual responses. Our results show that the traditional question of whether phases or magnitudes are more important for perception (Morgan et al., 1991; Oppenheim & Lim, 1981; Piotrowski & Campbell, 1982; Shapley et al., 1990) is inherently misled. These two entities are highly interdependent as only a special set of phases is consistent with any given valid collection of magnitudes, and thus, by encoding magnitudes, phase information is automatically and implicitly encoded as well.

The interdependence of phase and magnitude information has been noted before in the context of edge/bar detection. In their seminal paper, Morrone and Burr (1988) demonstrated that Gabor-like feature detectors are in phase at the location of lines and edges, and magnitudes [4] are peaked at these locations as well. In other words, while the magnitude information is insensitive to the *absolute* phase, it is correlated with the *relative* phase or phase relationships. Our results extend these previous findings by showing that the interdependence of (relative) phase and magnitude is a general rule, and the magnitude information in not only sufficient for edge/bar detection, but for object recognition in general.

It should also be pointed out that while the findings of some previous psychophysical studies have directly or indirectly suggested that it is the relative phase—and not the absolute phase—information which is important for object recognition (cf. Section 4.3), it has not been clear up until now whether a representation which ignores phase information at every sampling point (in space and frequency) would be able to preserve the relative phase information and encode scenes in such a way that they would be identifiable. Our results demonstrate that relative phase information *is* preserved to a degree which is sufficient for recognition. [5] It has been argued (Cavanagh, 1984) that "for shape recognition, the relative phase relations between frequency components as well as the amplitudes of the components are together sufficient to describe the shape" (p. 203). Our results show that the amplitudes alone would be sufficient for object description, as they implicitly encode relative phase information.

4.2. Invariances of complex cell representation

It is important to note that reconstructions vary in an interesting way. These variations reveal the invariance of the complex cell representation to contrast polarity of surfaces, and precise position of image structures such as edges and bars. These invariances can, in turn, subserve invariance to changes in illumination, background and deformation, as discussed below.

The main advantage of Gabor components over Fourier components is their locality, which makes them robust against variations such as changes in background, partial occlusion or slight distortion. At the same time, Gabor magnitudes retain to some extent one of the virtues of Fourier magnitudes, having local position invariance although within the confines of their Gaussian envelope. This insensitivity to the exact position of local image structures, as apparent in our reconstructions, provides robustness to object deformation and small scene distortions.

Our reconstructions also clearly show that a Gabor magnitude representation is ambiguous with respect to surface and edge contrast polarity. This attribute makes a complex cell representation invariant also to changes in background (Shapley & Gordon, 1985) which often cause the reversal of contrast polarity across object boundaries.

Insensitivity to contrast polarity is also important for invariance to changes in lighting. Surfaces and edges often change polarity at creases when illuminated from different directions. This is particularly important for polygonal objects. We investigated the stability of phase and magnitude information with respect to changes in the direction of lighting for stimuli composed of sharp creases (folded sheets at various orientations) and found that the Gabor magnitudes are significantly more stable than the corresponding phases. This result is highly intuitive since a change in the direction of lighting introduces a shift in the location of shadows, as well as reversal of contrast polarity (due to shadows), both factors to which Gabor magnitudes are invariant.

4.3. Behavioral evidence for the role of complex cells in object recognition

We have collected psychophysical data in a priming study using arbitrary images and their Gabor magnitude reconstructions. The data indicates that in spite of (potentially) strong differences in contrast polarity of image regions and exact line/edge position, an image and its reconstruction activate a common representation for object recognition. This finding suggests that the object recognition system is blind to image variations that complex cells are blind to. These results are consistent with those of previous psychophysical studies which have shown that priming is not dependent on contrast

---

[4] Our Gabor magnitudes correspond to what the authors referred to as "local energy." Our Gabor phases correspond to what they referred to as "arrival phase".

[5] Note that if all of the relative phase information were preserved, then we would always get one of the two reconstructions, either the exact replica of the original image, or its negative.

polarity of surfaces or edges, and even polarity reversals along a given edge (Subramaniam & Biederman, 1997), and others that have shown the contour detection and integration mechanisms are insensitive to the contrast polarity and spatial phase information (Field, Hayes, & Hess, 1993; Shapley & Gordon, 1985). Morrone and Burr (1988) have demonstrated that the detection of edges and lines in images can be performed based on the "local energy profile," which is quite similar to our Gabor magnitudes. This model has been shown to account for many perceptual effects such as Mach bands, the Craik-O'Brien illusion, and monocular rivalry (Burr & Morrone, 1990; Ross, Morrone, & Burr, 1989).

The consistency of the characteristics of a complex cell representation with those of the human recognition system, as revealed by our reconstruction and the various psychophysical results discussed above, suggests that complex cells may be what primarily underlie the representation mediating object recognition. On the other hand, it is likely that simple cells underlie representation for other processes such as shadow detection (Cavanagh & Leclerc, 1989), visuo-motor tasks, etc. in which the contrast polarity and/or exact position information is needed. The role of cells that cannot be classified as either simple or complex, and lie somewhere in between is more difficult to speculate about at this point. Depending on their degree of phase selectivity, these cells may be involved more in one of the two types of tasks—object recognition or (absolute) phase-dependent tasks.

5. Conclusion

It has been known for some time that individual complex cells encode magnitude information and are insensitive to spatial phase information. Our results show that while the phase information is lost at the level individual complex cells, the representation provided by a population of complex cells implicitly encodes the phase information that is needed for object recognition. While the relevant phase information (relevant phase relationships) are preserved in this representation, the partial loss of phase information (absolute phases) leads to some residual ambiguities. These ambiguities, however, are of a very interesting nature. They not only do not disrupt the recognition process, but rather, endow the recognition system with robustness to changes in illumination, background, and small distortions, i.e., changes to which human recognition system is highly robust.

In summary, our findings show that complex cells provide a representation which is highly advantageous in the discrimination-invariance trade off. They generalize over variations in background, lighting, and distortion while still being highly shape discriminating. These advantages may have been the basis for evolution's choice of complex cells as important component of higher vertebrate visual systems.

Acknowledgements

We thank Charles H. Anderson for very inspiring discussions from which arose the idea of reconstructing images from Gabor magnitudes by gradient descent. We also thank Ingo Rentschler, Rolf Würtz, Armand Tanguay, Tai Sing Lee, Yukiyasu Kamitani, and Stefan Schaal for their insightful comments. This work was supported by ARO's MURI Adaptive Optoelectronic Eyes and an ONR grant.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of Optical Society of America, 2*(2), 284–299.

Burr, D. C., & Morrone, M. C. (1990). Feature detection in biological and artificial visual systems. In C. Blakemore (Ed.), *Vision: coding and efficiency* (pp. 185–194). Cambridge: Cambridge University Press.

Cavanagh, P. (1984). Image transforms in the visual system. In P. Dodwell & T. Caelli (Eds.), *Figural synthesis* (pp. 185–218). New Jersey: Lawrence Erlbaum Associates.

Cavanagh, P., & Leclerc, Y. G. (1989). Shape from shadows. *Journal of Experimental Psychology: Human Perception and Performance, 15*(1), 3–27.

Daugman, J. (1993). Quadrature-phase simple-cell pairs are appropriately described in complex analytic form. *Journal of Optical Society of America, 10*(2), 375–377.

Daugman, J., & Downing, C. (1995). Gabor wavelets for statistical pattern recognition. In M. A. Arbib (Ed.), *The handbook of brain theory and neural networks* (pp. 414–420). Cambridge: MIT Press.

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, 2*(7).

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in Macaque visual cortex. *Vision Research, 22*, 545–559.

Dow, B. M. (1974). Functional classes of cells and their laminar distribution in monkey visual cortex. *Journal of Neurophysiology, 37*, 927–946.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of Optical Society of America, 4*(12), 2379–2394.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local association field. *Vision Research, 33*(2), 173–193.

Gallant, J. L., Braun, J., & Essen, D. C. (1993). Selectivity for polar, hyperbolic, and cartesian gratings in Macaque visual cortex. *Science, 259*, 100–103.

Grossmann, A., & Morlet, J. (1985). *Decomposition of functions into wavelets of constant shape, and related transforms, mathematics and physics. Lecture on recent results*. Singapore: World Scientific Publishing.

Hayes, M. H. (1982). The reconstruction of a multidimensional sequence from the phase or magnitude of its Fourier transform. *IEEE Transactions on Acoustics, Speech, and Signal Processing, ASSP, 30*(2), 140–154.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiolology, 195*, 215–243.

Jones, J., & Palmer, L. (1987). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology, 58*, 1233–1258.

Marcelja, S. (1980). Mathematical description of the responses of simple cortical cells. *Journal of the Optical Society of America, 70*(11), 1297–1300.

Morgan, M. J., Ross, J., & Hayes, A. (1991). The relative importance of local phase and local amplitude in patchwise image reconstruction. *Biological Cybernetics, 65*, 113–119.

Morrone, M. C., & Burr, D. C. (1988). Feature detection in human vision: a phase-dependent energy model. *Proceedings of the Royal Society London: Biological Sciences, 235*(1280), 221–245.

Movshon, J. A., Thompson, I. E., & Tolhurst, D. J. (1978). Receptive field organization of complex cells in the cat's striate cortex. *Journal of Physiology, 283*, 79–99.

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1997). Encoding of binocular disparity by complex cells in the cat's visual cortex. *Journal of Neurophysiology, 77*, 2879–2909.

Oppenheim, A. V., & Lim, J. S. (1981). The importance of phase in signals. *Proceedings of the IEEE, 69*(5), 529–541.

Palmer, S. E. (1999). *Vision science: photons to phenomenology*. Cambridge, Massachusetts: The MIT Press.

Piotrowski, L. N., & Campbell, F. W. (1982). A demonstration of the visual importance and flexibility of spatial-frequency amplitude and phase. *Perception, 11*, 337–346.

Poggio, G. F. (1972). Spatial properties of neurons in striate cortex of unanesthetized Macaque monkey. *Investigative Ophthalmology, 11*, 368–377.

Pollen, D. A., Gaska, J. P., & Jacobson, L. D. (1988). Responses of simple and complex cells to compund sine-wave gratings. *Vision Research, 28*, 25–39.

Pollen, D. A., & Ronner, S. F. (1981). Phase relationships between adjacent simple cells in the visual cortex. *Science, 212*, 1409–1411.

Pollen, D. A., & Ronner, S. F. (1982). Spatial computation performed by simple and complex cells in the visual cortex of the cat. *Vision Research, 22*, 101–118.

Ross, J., Morrone, M. C., & Burr, D. C. (1989). The conditions under which Mach bands are visible. *Vision Research, 29*(6), 699–715.

Schiller, P., Finlay, B., & Volman, S. (1976). Quantitative studies of single-cell properties in monkey striate cortex. I. Spatiotemporal organization of receptive fields. *Journal of Neurophysiology, 39*(6), 1288–1319.

Shapley, R., Caelli, T., Grossberg, S., Morgan, M., & Rentschler, I. (1990). Computational theories of visual perception. In L. Spillmann & J. S. Wener (Eds.), *Visual perception: the neurophysiological foundations* (pp. 417–447). San Diego: Academic Press.

Shapley, R., & Gordon, J. (1985). Nonlinearity in the perception of form. *Perception & Psychophysics, 37*(1), 84–88.

Skottun, B. C., De Valois, R. L., Grosof, D. H., Movshon, A., Albrecht, D. G., & Bonds, A. B. (1991). Classifying simple and complex cells on the basis of response modulation. *Vision Research, 31*(7/8), 1079–1086.

Spitzer, H., & Hochstein, S. (1988). Complex-cell receptive field models. *Progress in Neurobiology, 31*, 285–309.

Subramaniam, S., & Biederman, I. (1997). *Does contrast reversal affect object identification?* Paper presented at the Investigative Ophthalmology and Visual Science, Fort Lauderdale.