Generating complementary gray-level images for object recognition experiments using Gabor wavelet decomposition

I. KING and J. FISER
University of Southern California, Los Angeles, California

We present a software image processing methodology to generate complementary gray-scale images for visual perception experiments examining the interaction between the spatial localization and orientation sensitivity of visual perception. Specifically, our system is able to selectively mask or unmask specific resolutions, positions, and/or orientations of an input gray-level image using Gabor wavelet decomposition. The resulting representations are then combined to obtain partially reconstructed images for novel psychophysical experimentations.

The human visual system processes complex two-dimensional inputs that encode the environment. Although much research has probed the intimate relationship between the visual input and the neural representations and transformations of the visual system, most psychophysical experiments that investigate higher visual processes work usually with abstract visual inputs, such as line drawings, symbols, or letters. Furthermore, experiments utilizing gray-scale images typically use simple visual stimuli, such as Gaussian blobs, gratings, and plaid patterns, to test low-level issues (e.g., contrast sensitivity).

There are two main reasons why gray-scale images are not used more extensively in higher level psychophysical experiments. First, it is difficult to manipulate gray-scale images in a controlled way, which is required for generating good stimuli. Second, it is not clear what should be manipulated in order to ask meaningful questions about higher visual information processing.

Our objective is to present a software methodology that can be used for controlled gray-scale image manipulation and that has strong relations to known facts about biological visual information processing, making it suitable to pursue issues of higher level vision. Using this methodology, we plan to map out what is important and what is less relevant in particular visual object recognition tasks.

Rationale of the Methodology

Biederman and colleagues have investigated high-level issues of form perception with line-drawing experiments (Biederman & Cooper, 1991, 1992; Biederman & Gerhardstein, 1993; Biederman & Ju, 1988). They used different object recognition tasks and analyzed the effect of visual priming by reaction time and error rate measurements. The conclusion of these studies was that human object recognition proceeds via intermediate representations that use part-based and viewpoint invariant descriptors.

In order to develop further our understanding about object recognition, one needs to work with more realistic inputs (e.g., gray-scale images). We selected one particular experiment, priming with complementary images, to demonstrate why this transition is necessary, and how it can be done by using our methodology.

Complementary-image object recognition tasks are known to suggest the importance of intermediate representations in visual object recognition (Biederman & Cooper, 1991). The stimuli in these tasks are created by deleting a subset of image features (e.g., edges and vertices) from line drawings of objects in order to create two complementary images of each object. Although complementary images share no edges and vertices, they are created in such a way that different components of the object could be recovered from the existing segments of either complementary image (see Figure 1).

More formally, let a binary image A be an $m \times n$ array with elements $a_{i,j} \in \{0,1\}$, $1 \leq i \leq m$, $1 \leq j \leq n$, where $a_{i,j}$ is the pixel value of the binary image A at position (i,j). Here, the binary image A contains a complete line drawing of an object. Complementary images $\{A_1, A_2\}$ are two binary images with the property that (1) $A = A_1 \cup A_2$, and (2) the intersection of the complementary images $A_1 \cap A_2 = 0$, where 0 is a specialized binary image, with $a_{i,j} = 0$, $\forall i,j$.
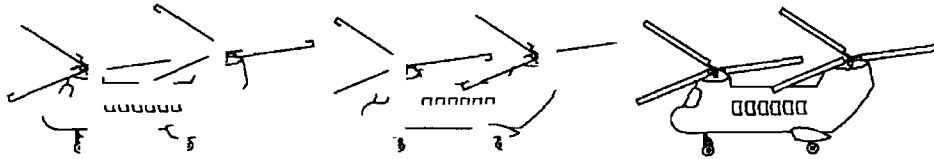
Figure 1. Complementary images of a helicopter in (a) and (b). The union of these two binary complementary images is the complete image of a helicopter shown in (c).

The typical process of generating these binary images is to perform line extraction from a sampled gray-scale image by retaining high contrast intensity edges. The complementary images are then generated by dividing the full image into two approximately equal binary complementary images through contour deletion from each complement of the object, as shown in Figure 1. The partitioning of these complementary pairs of images is done in such a way that (1) $\Sigma_i \Sigma_j a_{i,j} \in A_1 \approx \Sigma_i \Sigma_j a_{i,j} \in A_2$, the pixel counts in A_1 and A_2 are almost identical, and (2) long lines are preserved as much as possible to eliminate dashed lines (alternating line segments), preserving maximum connectivity with fewest line segments.

The results of the experiments with these images showed that the same amount of priming existed whether an element of a complementary pair was primed by itself or by the other complementary image (Biederman & Cooper, 1991). Moreover, the amount of priming was the same irrespective of whether the priming picture was a complementary or an intact version of an object. Hence, Biederman and Cooper concluded that priming is mediated not by particular edges and vertices present in an object but rather by meaningful components or parts of the object. This suggests that humans recognize objects by first recognizing the components belonging to the object.

The above investigation was conducted on images that are based on line drawings composed of easy-to-mask-out lines and, consequently, had the following shortcomings:

1. These contour images are idealized representations of real-world objects. Typically, line drawings are formed by tracing the essential contour of intensity boundaries of an object. Here, *essential* means retaining only the lines that convey information about parts. Such a clean representation is not common in real situations. Furthermore, it is a subjective task to select the particular line segments. Therefore, one might wonder whether the results hold only for the specific stimuli used in the experiment or capture general characteristics of object recognition.

2. Deleting a line segment means introducing extra information by line ends at the border of deleted areas. The importance of line ends in object recognition is well documented (Kennedy, 1988; Lesher & Mingolla, 1993) for completion of illusory contours. It becomes a question of whether the particular process of equal priming is due to the information kept about the objects, or, more to the information about the line ends, which is similar in both complementary images (i.e., at corresponding cuts, their positions and their orientations are the same).

3. Even if the issue of line ends is neglected, two images that are complementary on the pixel level can be identical in a different representation. Eventually, they do become identical according to the results demonstrated in the experiments (Biederman & Cooper, 1991). In order to follow the visual process "upstream" and to gain more insight about the set of descriptors by which this similarity is captured, one must go beyond pixel representation. Pixels are suitable to model representation at the photoreceptor level but not for representations in higher cortical areas.

In order to produce more suitable visual stimuli, information about representing visual input in biological systems must be incorporated. Growing evidence suggests that the human, as well as other mammalian, visual systems have a localized multiresolution and multiorientational representation that forms one of the earliest cortical representations of the visual scene (De Valois & De Valois, 1988; Kulikowski, Marčelja, & Bishop, 1982; Marčelja, 1980). Neurophysiological recordings in experiments with bars and gratings have demonstrated that the majority of the neurons involved in this representation are sensitive to oriented intensity gradients within a narrow range of spatial frequency restricted to a small area of the visual scene (Hubel & Wiesel, 1959, 1962). The characteristics of these cells are reminiscent of spatially localized, oriented bandpass filters and can be modeled well by Morlet wavelets generated with Gabor functions (Daugman, 1985; Jones & Palmer, 1987; Marčelja, 1980). Gabor functions are known to be optimal in the sense of retaining the maximum combined local position and local spatial frequency information within a localized area (Gabor, 1946). Therefore, a multiscale multiorientation Morlet wavelet representation of the visual scene is not only a good approximation of the early representation mentioned above but also an efficient data-encoding method (Mallat, 1989).

Little is known about other types of representations in the visual pathways, and most of those representations involve other dimensions, such as disparity, motion, or color information. If gray-scale still pictures are considered to be inputs, the Morlet wavelet representation could be, in theory, the dominant base for the subsequent processes and representations in object recognition. If this is so, manipulation of the content of the visual input should be diagnostic as to what type of more complex representations are derived from this early one in form perception.

For this reason, this paper proposes to represent 256-level gray-scale images with a four-dimensional array of Gabor function coefficients (hereafter referred to as *Gabor rep-*

resentation). Two dimensions represent spatial location on the visual scene. The third and fourth dimensions represent the scale, s, and orientation, o, indices respectively. A coefficient in position (x,y,s,o) describes the value after convolving the original image with a Gabor function with the specific (s,o) pair at the spatial location (x,y).

It can be shown mathematically that, by a linear inverse transformation, the original image can be reconstructed from this representation (Daugman, 1988, 1989). Moreover, it is suitable to handle the shortcomings of line drawings listed above in the following way. First, the final outcome of any manipulation on this representation if inverted back will give a gray-scale image much like the original one. Second, deleting any coefficient from the representation will mean removing the energy of the specific oriented gray-scale information on the defined spatial scale at the given location without introducing any extra information to the original image. Note that the flexibility offered by this method is impossible to achieve by Fourier transformation, because the Gabor transformation has two extra parameters that can be controlled independently (i.e., location and orientation), which the global bandpass–filtering method cannot control at all. Third, due to the similarity of this representation to some aspect of the early cortical representation in humans, it can be used for carrying out psychophysical experiments on humans and animals and on computer models in parallel (Fiser, Biederman, & Cooper, 1994). In this way, one can gain more insights into what aspects of human vision this representation fails to capture.

In the following sections, the formalism of Gabor decomposition that leads to a method to generate complementary images will first be presented. Finally, some results and further developing directions will be discussed.

Background of Gabor Wavelet Decomposition

A signal varying with time or, in our case, a quantized intensity array is decomposable using Fourier transformation into elementary components of sine and cosine functions that make up the signal. Similarly, the Gabor wavelet transform can be used to decompose the image into the spatial and frequency domain of an input image using elementary Gabor functions (Gabor, 1946). The function itself is a complex sinusoid that controls the orientation enveloped under a Gaussian window, which controls the spatial localization specifying an area of interest. The two-dimensional Gabor function and its Fourier transform are defined by Daugman (1985) as

$$g(x,y;u_0,v_0;\sigma_x,\sigma_y) = e^{-\left(\frac{x^2}{2\sigma_x^2}+\frac{y^2}{2\sigma_y^2}\right)}e^{2\pi i(u_0 x+v_0 y)}, \quad (1)$$

and

$$G(u,v;u_0,v_0;\sigma_x\sigma_y) = e^{-2\pi^2(\sigma_x^2(u-u_0)^2+\sigma_y^2(v-v_0)^2)}. \quad (2)$$

The first exponential term in Equation 1 is the Gaussian envelope that dictates the spatial width of the localization by the parameters σ_x and σ_y. The second term in Equation 1 is the complex sinusoid modulating the direction of the spatial orientation; its modulation is controlled
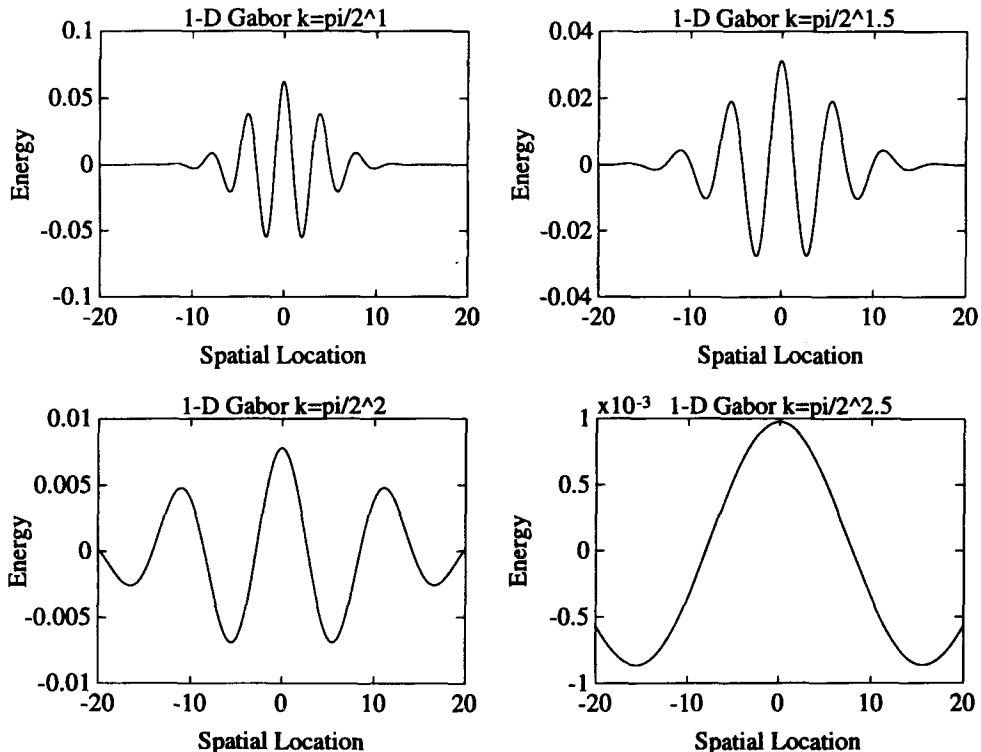

Figure 2. One-dimensional characterization of an even-symmetric Gabor filter at four different resolutions with $\upsilon = 0,1,\ldots,3$ (the order is from left to right, top to bottom).

by the u_0 and v_0 parameters.[1] The larger the value, the more ripples will be generated due to the change in frequency. The angle is calculated by $\theta = \arctan(v_0/u_0)$. An even-symmetrical (cosine) one-dimensional Gabor filter is shown in Figure 2. The cosine function is an even function with respect to the origin and is sensitive to lines, whereas the odd-symmetrical (sine) filters are more suitable for detecting intensity step edges in images (Mehrotra, Namuduri, & Ranganathan, 1992).

Decomposition of an image (an intensity function or a gray level distribution), $I(\vec{x})$, using a Gabor wavelet, is defined as

$$I^o_{\upsilon v}(\vec{x}) := I_{\upsilon v}(\vec{x}) * \Psi_{\upsilon v}(\vec{x}), \tag{3}$$

where $I^o_{\upsilon v}(\vec{x})$ is the resulting image, $\Psi_{\upsilon v}(\vec{x})$ is a Gabor wavelet kernel, υ and v are indices to the scale and orientation of the Gabor kernel function, and $*$ is the convolution operator.[2] Using five scales and eight orientations, the kernels are defined by the equation

$$\Psi_{\upsilon v} = \frac{\pi}{2^{1+(\mu/2)}} e^{i(\pi v)/8}, \upsilon \in \{0,\dots,4\}, v = \{0,\dots,7\}. \tag{4}$$

Hence, there are $5 \times 8 = 40$ decomposed complex image pairs (80 decomposed pictures in total) after the convolution operation. The orientation of these filters increases in a $22.5°$ increment covering from 0 to π. The range from π to 2π is symmetrical; therefore, we need not generate filters to cover that specific range (for further information, see Doursat et al., 1993, and von der Malsburg & von Seelen, 1992).

The coverage of the frequency space by the five scales of Gabor filters is demonstrated in Figure 3 (top). The spac-

ings in the frequency domain of the filters are measured in octaves. The octave is defined by the measurement between two peak frequencies as $\log f_1/f_2$. Thus, the octave in Figure 3 (top) has a constant increment of 0.5. In other words, the peak frequency for each Gabor filter is doubled, relative to the preceding filter, to ensure an adequate logarithmic coverage of the Fourier frequency space. The summation of the coverage is shown in Figure 3 (bottom).

The convolution operation is computationally intensive when calculating the product of the image with the Gabor masks. To reduce the computation time, we use the convolution theorem, $\mathcal{F}(f * g) = \sqrt{2\pi}\mathcal{F}(f)\,\mathcal{F}(g)$, to calculate the convolution of the image with the filter [$\mathcal{F}(f)$ stands for the Fourier transform of function f]. According to the theorem, the result of the convolution is equivalent to the result of an inverse Fourier transform of the product of the Fourier-transformed input and the Fourier-transformed filter. Moreover, since the masks typically need not be modified once a set of satisfactory parameters are found, the Fourier-transformed Gabor filters (kernels) can be precomputed to reduce the execution time.

In the reconstruction process, each decomposed image is first convolved with its corresponding kernel (i.e., with the filter of that orientation and scale by which the decomposed image was obtained). The same convolution theorem is applicable in this stage as well. Since a complex Gabor is the same as its complex conjugate, the two convolution operations involve exactly the same steps with the same kernel functions. Next, the resulting images are subjected to pointwise summation to reconstruct the real part for the final output image and the imaginary part, which is summed to approximately a constant zero image.
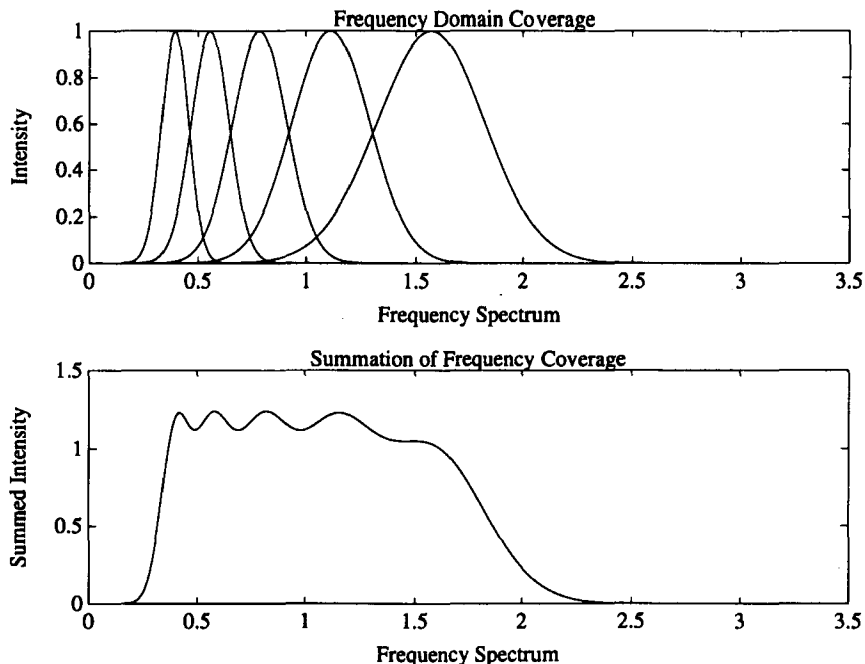

Figure 3. (top) Coverage of the frequency space $\upsilon = 0,1,\dots,4$ by wavelets of different resolution levels. (bottom) The sum of the five frequency bands and their coverage resulting from the top image.

In this reconstruction process, different manipulations of the information content of the representation are possible. Since this representation expresses the oriented spatial frequency content of each small area of the original image exactly, it is easy to produce an image that has only, say, vertical spatial frequencies in two separate octaves and no other components in the entire image by selecting the proper scale and orientation at specific image locations.

One important characteristic of the Gabor functions is that they form a complete but nonorthogonal basis set. Nonetheless, a full and complete reconstruction of the decomposed images preserving all the fidelity is theoretically possible (Daugman, 1988, 1989; Porat & Zeevi, 1988). In reality, this reconstruction procedure is hampered by quantization, calculation, and round-off errors. Therefore, the reconstructed image loses its highest and lowest (DC) frequencies. The nonorthogonality of the Gabor filters also means that the coverages of these filters in the frequency domain overlap with each other due to the Gaussian-like envelope functions. This overlapping creates regions of emphasis in the frequency domain, as illustrated in the multimodal plot in Figure 3 (bottom), which further degrades the reconstruction process. Furthermore, in reality, the ends of the spectrum also lose their fidelity due to the gradual decrease of the function. Nonetheless, the Gabor functions are suitable for visual perception experiments, where only fair quality of reconstruction is required.

System Environment

A functional diagram of the system is illustrated in Figure 4. Our system first decomposes the input gray-level image into 80 (40 pairs of real and imaginary images) Gabor-filtered images and then constructs the final image by selecting the desired scale, orientation, and the specified locations in the image for visual experimentation.

The selection process for manipulating the scales and orientations of the decomposed images is made with a control file. The control file is a text file with a matrix table indicating which spatial scale and orientation images are to be included in the summation process. An entry of a "0" or "1" controls the exclusion or the inclusion, respectively, of a particular decomposed image. To specify the location of points, another control file contains a list of x and y coordinate locations for the inclusion process. These locations can be selected in a graphical interactive way by clicking on different positions in the image. It is possible to use the orientation and spatial frequency information of (1) only these points, (2) all but these points, or (3) all points around these points within a fixed window size in the reconstruction process. A sample of the two files is shown in Figure 5.

Results and Applications

To generate the decomposed images, we use a combination of C programs and MATLAB M-files on Unix workstations (both SUN and DEC workstations). The original and its fully reconstructed image are illustrated in Figure 6. A small set of Gabor filters are shown in Figure 7. The images are arranged in the order from left-to-right and top-to-bottom fashion showing five different scales and five different orientations, $v = v = \{0,1,...,4\}$.

To demonstrate the features of our system, we present results with a typical real-world image. The set of images containing different scale (resolution) information with all orientations integrated is shown in Figure 8. Each image is composed of all eight orientations within the same scale. Adding up these five images produces Figure 6 (right).

It is clear that if all the coefficients in a localized area around (x,y) spatial position in the Gabor representation are set to zero, after the inversion, the same localized area will be blank in the two-dimensional image. This is true if
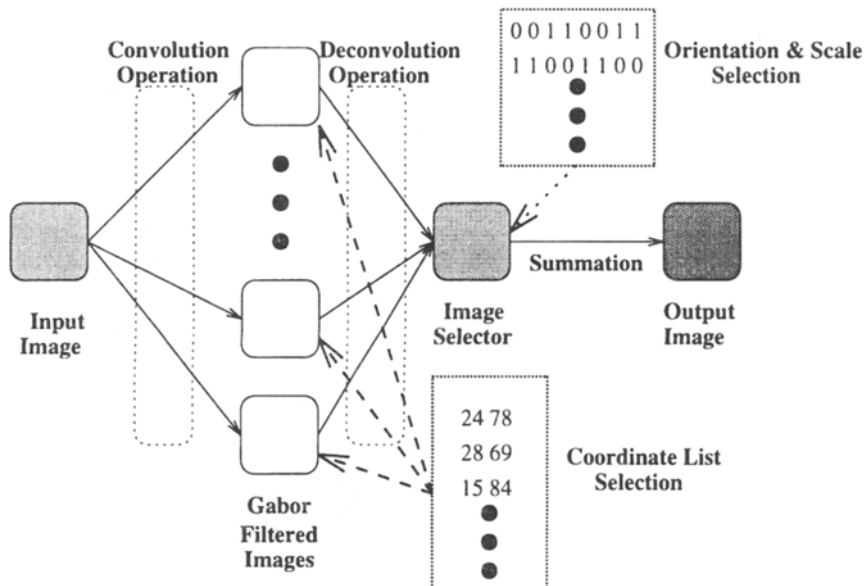
Figure 4. A system schematic of Gabor banks for decomposition of the input image and production of the final output image.

```
1  1  1  1  1  1  1  1        20,4,1

  0  0  0  0  0  0  0  1        26,32

  0  0  0  0  0  0  0  1        37,12

  0  0  0  0  0  0  0  1        25,36

  0  0  0  0  0  0  0  1        28,37

                                 ⋮

        (A)                    (B)
```

Figure 5. (A) A sample control file for selecting the orientation (= 8) and scale (= 5) images. The file selects all orientations in the largest scale and all the scales with the orientation of 157.5°. (B) A sample control file for the spatial coordinates that will be included in the summing process. The first line is a control line. The first number (= 20) indicates how many points are present in the file. The second number (= 4) is the window size used during the reconstruction processing. The third number (= 1) is a flag that tells the program whether to include (= 1) or exclude (= 2) the ensuing coordinates as the input or their complement.

the area is sufficiently large so that the overlap effect from neighboring positions is negligible.

The method is suitable to delete the visual content of an area of arbitrary shape in the gray-scale image without introducing artifacts by setting the coefficients to zero. This is useful in experiments where localized parts of the presented objects used as stimuli have to be deleted.

The essential question in the complementary-image priming experiment described earlier was how different the two complementary versions of an object picture are from the point of view of human object recognition measured by the priming effect. The results suggested that, despite the fact that the two complementary images carried no common information on pixel level, they were perceived as identical. Here, information meant a pixel value turned to black, and white pixels meant background (i.e., no information).

The generalization of the question above to gray-scale images can be given as follows. How different will two complementary images of an object be perceived by hu-

mans (measured by priming) if the images carry no common gray-scale information? Here, no common information means that the two complementary images give (1) precise information about nonoverlapping regions of the original picture, (2) no information about the regions of the other complementary image, and (3) information of smoothly degrading quality along any line starting in the center of one of its own regions and running to the center of a neighboring region belonging to the other complementary image. Precise information means equivalent to the fully reconstructed image. Depending on the quality of the reconstruction, this might have substantially less dynamic range than the original image (see Figure 6). Therefore, in the experiments, the baseline effect must be measured with the fully reconstructed version rather than using the original image.

Complementary images of the kind described above can be produced in the following way with the system. First, the original image is decomposed into the Gabor representation. In the case of a 128×128 size image this means $128 \times 128 \times 40 \times 2$ coefficients. Next, a relatively small subset of the 128×128 positions ($< 1,000$) are selected in the following way. First, 10–20 "anchor points" are manually selected using the mouse so that the anchor points are evenly distributed in the region occupied by the object. These anchor points are typically salient points representing areas of interest (e.g., intensity changes, high curvatures, etc.). Next, a window is defined so that all positions within the window around each anchor point are included in the set of selected positions. For example, if a window size of 8 is chosen, the closest 64 positions around each anchor point will be included, which gives less than 1,280 points if 20 anchor points were selected, assuming no significant overlapping occurs among the windows.

Next, all coefficients in the Gabor representation that belong to nonselected positions are set to zero. In other words, all information based on the selected points is kept, and all the rest is neglected. The information in the selected positions is enough to produce a clearly recognizable view of the object after the inverse Gabor transformation. Finally, the anchor points are separated into two equal groups by the experimenter such that both groups cover the entire

Figure 6. (left) The lena image with 128×128 pixels with 8 bits (256) gray levels obtained from an original 512×512 standard lena image by direct 4:1 subsampling. (right) The reconstructed image by summing up all scales and all orientations.
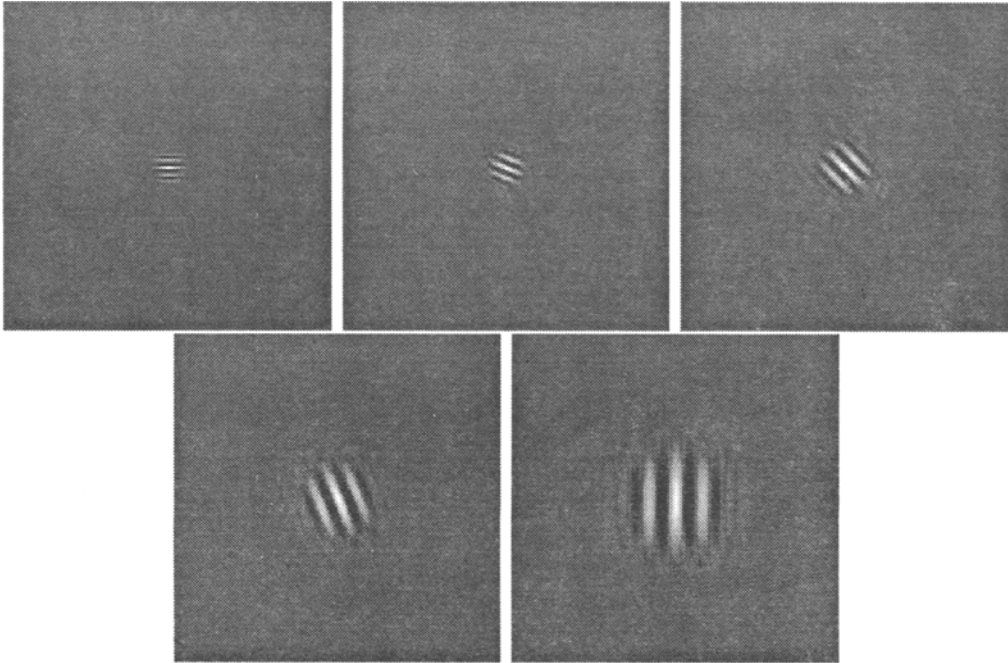
Figure 7. Cosine Gabor filters. Here are five different scales of Gabor filters with five different orientations (from 0° [top-left] to 90° [bottom-right] with an increment of 22.5°).
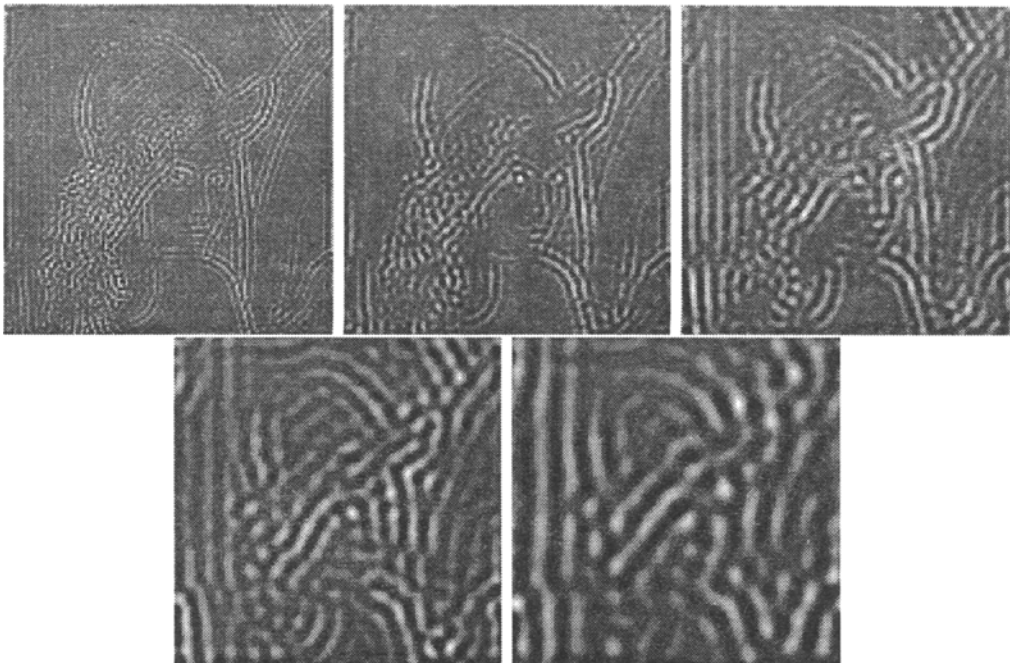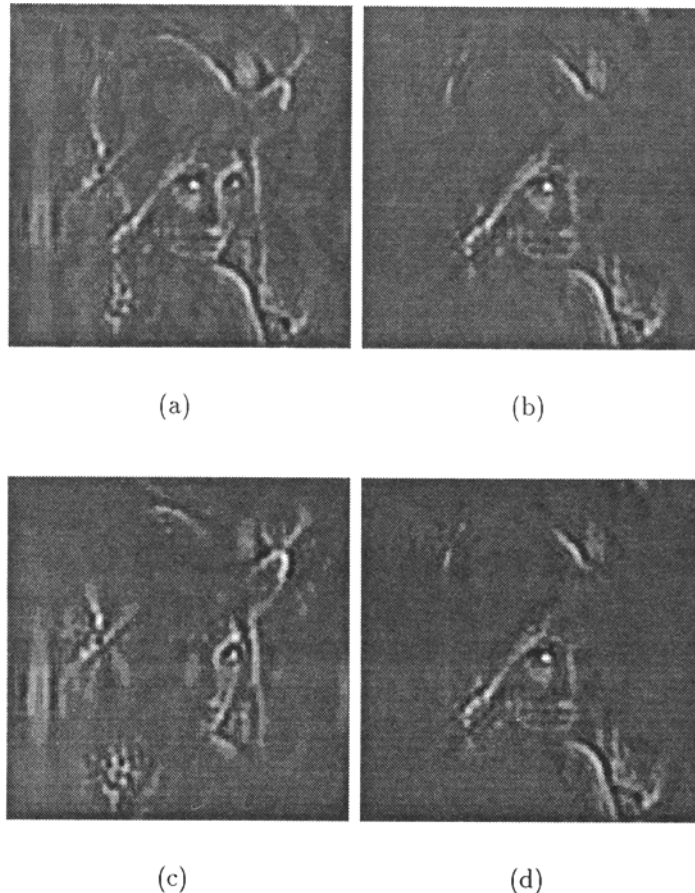

Figure 8. Different scales $\mu = 0, \ldots, 4$ with all orientations. At the lower scale (with smaller spatial areas), Gabor filters act like high-pass filters, whereas at the higher scale (with bigger spatial areas), the Gabor filters act like low-pass filters. The summation of these five images will result in the reconstructed image shown in Figure 6 (right).

Figure 9. Generating complementary images with the partial reconstruction method. The full reconstruction using information of all the anchor points in all 14 clusters is shown in (a). In (c) and (d), the two complementary images of (a) are shown using 7-7 clusters which were selected alternatively, thereby containing detailed information about different parts of the original image. In (b), the resulting image by subtracting image (c) from image (a) is shown.

object and they intermingle evenly. The two final sets of positions include the two sets of anchor points and their surrounding positions within the window. In this way, two Gabor representations are obtained: one for each set of positions with all the coefficients belonging to their positions and zeros in the complementary positions and in the neglected positions. If the anchor points are spaced sufficiently sparsely, inverting back the two representations separately will produce two complementary images with the desired characteristics.

Figure 9 demonstrates how the system can be used for creating complementary images with the method described above. This reconstruction used 14 clusters (7-7 for the two complementary images), each cluster having 50–55 anchor points in the image. Information on all five scales and eight orientations in these points was summed up.

Discussion

Although complementary gray-scale images can be created in many different ways, the method described above is the closest reproduction of the original line-drawing ex-

periment by Biederman and Cooper (1991), in that only the dimension of spatial position was used to create the two complementary images, without making any modification in the other dimensions. However, the Gabor representation allows more flexibility and control over spatial scale and orientation information in each position.

The simplest manipulation is essentially band-pass filtering when the spatial positions are the same in the two complementary images; however, assuming six spatial scales, one complementary image uses the coefficients of, say, Scales 1, 3, and 5, while the other uses Scales 2, 4, and 6. This type of complementary images can be obtained by the traditional Fourier transformation as well. However, with the Gabor representation, one can mix spatial position and spatial frequency alternation by separating the entire set of positions into intermingled subsets, so that neighboring positions will use information from different scales. This will produce images that are suitable for testing how far apart the scales of information and their anchor positions can be before the human visual system fails to extract information needed for further processing.

A further possibility is to involve the dimension of orientation. Many researchers have reported specific interactions between different orientation and scale channels (e.g., Thomas, Olzak, & Shimozaki, 1993). With our method, it becomes possible to investigate how these interactions manifest themselves in higher level object recognition, since this method is suitable to test the effects of such local interactions in the visual system, whereas approaches based on the traditional global Fourier transformation are not.

Due to the Gabor filter's biological relevance, many researchers are modeling the primary visual cortex using this multiresolution approach. Research in visual perception functions, such as texture analysis (Jain & Farrokhnia, 1991) and boundary perception (Manjunath & Chellappa, 1993), is also conducted using these decomposition filters. Practical application, such as the face recognition system, also utilizes the Gabor decomposition as its preprocessing module (Buhmann, Lange, von der Malsburg, Vorbruggen, & Wurtz, 1992; Lades, Vorbruggen, & Buhmann, 1993). Our system can be used in these areas for testing particular hypotheses about the representation by Gabor filters.

Conclusion

In this paper, we have demonstrated a method to decompose a gray-scale image using the Gabor wavelet transformation in order to mask out specific orientation, scale, and spatial position. The resulting images are manipulated in a controlled way on the basis of their spatial frequency, orientation selectivity, and locality. The obtained representation allows us to create experimental images with novel characteristics to test different theories of visual perception.

REFERENCES

BIEDERMAN, I., & COOPER, E. E. (1991). Priming contour-deleted images: Evidence for intermediate representations in visual object recognition. *Cognitive Psychology, 23,* 393-419.

BIEDERMAN, I., & COOPER, E. E. (1992). Size invariance in visual object priming performance. *Journal of Experimental Psychology: Human Perception & Performance, 18,* 121-133.

BIEDERMAN, I., & GERHARDSTEIN, P. C. (1993). Recognizing depth-rotated objects-evidence and conditions for 3-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception & Performance, 19,* 1162-1182.

BIEDERMAN, I., & JU, G. (1988). Surface versus edge-base determinants of visual recognition. *Cognitive Psychology, 20,* 38-64.

BUHMANN, J., LANGE, J., VON DER MALSBURG, C., VORBRUGGEN, J., & WURTZ, R. (1992). Object recognition with Gabor functions in the dynamic link architecture. In B. Kosko (Ed.), *Neural networks for signal processing* (pp. 121-159). Englewood Cliffs, NJ: Prentice Hall.

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.

DAUGMAN, J. G. (1988). Complete discrete 2-D Gabor transforms by

neural networks for image analysis and compression. *IEEE Transactions on Acoustic, Speech, & Signal Processing, 36,* 1169-1179.

DAUGMAN, J. G. (1989). *Non-orthogonal wavelet representations in relaxation networks: Image encoding and analysis with biological visual primitives.* New York: Adam Hilger.

DE VALOIS, R. L., & DE VALOIS, K. K. (1988). *Spatial vision* (2nd ed.). New York: Oxford University Press.

DOURSAT, R., KONEN, W., LADES, M., VON DER MALSBURG, C., WORBRÜGGEN, J., WISKOTT, L., & WÜRTZ, R. (1993). *Neural mechanisms of elastic pattern matching* (Tech. Rep. IR-INI 93-01). Bochum: Ruhr Universität, Institut für Neuroinformatik .

FISER, J., BIEDERMAN, I., & COOPER, E. E. (1994). Are the direct outputs of Gabor filters sufficient for human object recognition or are they only the prior stage for intermediate representations [Abstract]? *Investigative Ophthalmology & Visual Science, 35,* 1625.

GABOR, D. (1946). Theory of communication. *Journal of the Institution of Electrical Engineers, 93,* 429-457.

HUBEL, D. H., & WIESEL, T. N. (1959). Receptive fields of single neurons in the cat's visual cortex. *Journal of Physiology, 148,* 574-591.

HUBEL, D. H., & WIESEL, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology, 160,* 106-154.

JAIN, A. K., & FARROKHNIA, F. (1991). Unsupervised texture segmentation using Gabor filters. *Pattern Recognition, 24,* 1167-1186.

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

KENNEDY, J. M. (1988). Line endings and subjective contours. *Spatial Vision, 3,* 151-158.

KULIKOWSKI, J. J., MARČELJA, S., & BISHOP, P. O. (1982). Theory of spatial position and spatial frequency relations in the receptive fields of simple cells in the visual cortex. *Biological Cybernetics, 43,* 187-198.

LADES, M., VORBRUGGEN, J. C., & BUHMANN, J. (1993). Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers, 42,* 300-311.

LESHER, G. W., & MINGOLLA, E. (1993). The role of edges and line-ends in illusory contour formation. *Vision Research, 33,* 2253-2270.

MALLAT, S. G. (1989). Multifrequency channel decompositions of images and wavelet models. *IEEE Transactions on Acoustics, Speech, & Signal Processing, 37,* 2091-2110.

MANJUNATH, B. S., & CHELLAPPA, R. (1993). A unified approach to boundary perception: Edges, textures, and illusory contours. *IEEE Transactions on Neural Networks, 4,* 96-107.

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

MEHROTRA, R., NAMUDURI, R., & RANGANATHAN, N. (1992). Gabor filter-based edge detection. *Pattern Recognition, 25,* 1479-1494.

PORAT, M., & ZEEVI, Y. Y. (1988). The generalized Gabor scheme of image representation in biological and machine vision. *IEEE Transactions on Pattern Analysis & Machine Intelligence, 10,* 452-468.

THOMAS, J. P., OLZAK, L. A., & SHIMOZAKI, S. S. (1993). The role of Fourier components in discrimination between 2 types of plaid patterns. *Vision Research, 33,* 1573-1579.

VON DER MALSBURG, C., & VON SEELEN, W. (1992). *Annual report 1992.* Bochum: Ruhr Universität, Institut für Neuroinformatik.

NOTES

1. It obeys the Euler equation, $e^{\pm i\theta} = \cos(\theta) \pm i \sin(\theta)$.
2. It is typically defined as $h(t) = \int_0^t f(\tau)g(t-\tau)d\tau$.