Cortical Representation of Texture Primitives

A.E.J. Walford and M.E. Jernigan
Department of Systems Design Engineering
University of Waterloo
Waterloo, Ontario, Canada N2L 3G1

Abstract

The apparent spatial frequency and orientation organization of the primary visual cortex is used as a basis for texture description. A Frequency, Orientation, neural firing Rate, and spatial Phase (FORP) representation is proposed for the analysis of natural textures and the synthesis of test textures. Higher order texture analysis such as discrimination and segmentation in this FORP space is discussed.

Keywords: texture analysis, vision models, primary visual cortex models

Introduction

Many researchers have tried to duplicate the ability of the human visual system to segment natural scenes based on the different texturing of surfaces. We choose to define texture as that property of surfaces that can be described by the local pattern of spatial variation of intensity. We also take as different textures those that can be differentiated by human perception. In this light, we base our analysis of texture on a model of the early human visual system. In particular, we choose a model of the primary visual cortex since both orientation and spatial frequency information exist here. [1]

We use this information in an orthogonal feature extraction space to generate simple test textures and analyse natural textures.

Texture and the Visual Cortex

Given that the primary visual cortex is well suited to texture computations, is there any evidence that it actually performs them? The following evidence indicates that this could indeed be true.

First of all, Julesz[2] mentions the high speed discrimination ability of human subjects. This implies that texture discrimination is a low level visual function and must be early in the visual chain. Kimchi and Palmer,[3] through the use of perceptual experiments found that texture is processed separately in the visual system from shape and structure. Lastly, Berlucchi and Sprague[4] use lesion studies to deduce that shape and structure encoding does not exist in the primary visual cortex. They suggest the primary visual cortex could be used for texture analysis.

The primary visual cortex appears well suited for texture analysis, it appears to actually perform texture computations, and it can use this information to improve image segmentation.

A Visual Cortex Model

Pollen and Ronner[1] describe a model of the primary visual cortex which outlines various functions of cortical neurons. These functions include the retinotopic spatial map, ocular dominance, orientation, spatial frequency and spatial phase. Hubel and Wiesel[5] first described a small separate processing region in the primary visual cortex as the *hypercolumn*. The *hypercolumn* was responsible for analysing a small area of the visual field for orientation information. The hypercolumn has become the name for the 0.5 mm wide cortical area that contains orientation and frequency selectivity neurons for a small visual region. There exist hypercolumn regions to cover all of the visual field.

Our model of the primary visual cortex is based on this hypercolumn structure. One hypercolumn is modeled as a three dimensional space. The space consists of a spatial frequency axis, an orientation axis, and a spatial phase axis. Each point in this space corresponds to a neuron that is selective to a particular set of frequency, orientation, and phase. A magnitude component is also included in this space to account for the strength of response to this particular set. The magnitude is coded by the neuron as a neural firing rate. This model is a modified two dimensional Fourier space with all symmetric regions removed. The model has been named **FORP** (spatial Frequency, Orientation, neural firing Rate, and spatial Phase).

The FORP based tools that were developed allowed us to both synthesize and analyse textures. One of the tools allowed a window of a texture image to be chosen and be represented in FORP space. The other allowed the experimenter to place individual points or group of points in the space and then have the texture generated that corresponded to this pattern of firing neurons. Figures 1 thru 6 are outputs of these programs.

Figure 1 shows a representation of FORP space and the corresponding synthesized texture. It shows the three axes of frequency (range of 0 to $F_S/2$), orientation (range of $0°$ to $180°$), and neural firing rate (normalized to a range of 0 to 1). There is a grid on the frequency/orientation plane, and the intersection of two grid lines corresponds to a neuron or a set of neurons. The intensity of the bar at a frequency/orientation point represents the phase. The lowest intensity corresponds to a phase of 0 and the highest to a phase of 2π. If there were a set of phase sensitive neurons then the bar's intensity would represent which phase neuron in this set was responding maximally. The image beside the graph contains a homogeneous texture generated from the FORP space data. The little square in the bottom left corner of the image delineates the local region that is represented by the FORP space. The remainder of the texture image is a mosaic-like repetition of this small subimage.

To demonstrate the nature of FORP space we have performed a number of simple texture syntheses. In Figure 1, a single point is placed at a frequency of $F_S/16$ (that is 1/16th of the spatial sampling frequency), and an orientation of $110°$ (note that $0°$ is vertical). The resulting synthesized texture is a spatial sinusoid at $110°$ and has 2 periods in its 32 by 32 sub-window. This point is moved along the frequency and orientation axes (Figure 2). Note that the frequency change modifies the sinusoidal spacing and the orientation change modifies the sinusoid's angle.

To answer the question, "What will natural textures look like in FORP space? ", we performed some texture analyses shown in Figures 3 thru 6. Images of the natural textures oriental rattan,[6] diatom,[7] and J-Cloth™ were each sampled with windows at different positions. The window size was 64 by 64 taken out of a 256 by 256 image, except the diatom samples which were done with a 32 by 32 window. The most important aspect of these FORP representations is their similarity in shape for the same texture, and their dissimilarity for different textures. In Figures 3 and 4, both FORP space representations of oriental rattan look quite similar. The majority of neural response is along the $0°$, $90°$ and $180°$ lines. This is understandable when one realizes that oriental rattan is composed mostly of horizontal and vertical edges. In Figure 5 the FORP space of diatom

looks quite different from that of rattan. The neural responses are spread quite evenly in orientation. This is to be expected of a texture that is composed of circles, since a circle has edges of all orientations. Again, in Figure 6 the FORP space of J-Cloth™ is quite different from previous ones.

D'Astous[8] comments on this similarity of frequency domain representations and its importance: "... the power spectrum is fairly invariant to minor changes in structure caused by either low magnitude additive noise, or by small deviations in the periodicity of the texture. This is of particular relevance to the problem of discriminating natural textures which tend to be noisy and, though many textures exhibit regularity to a certain extent, are not strictly periodic."†

Required Accuracy

The accuracy of the FORP parameters was studied to see how it affected the representation of textures. A window of a texture was converted to FORP space at various frequency and orientation accuracies. This representation was then used to generate a texture which was compared to the original.

The amount of orientation information required depended on the type of texture. A synthetic texture with only horizontal and vertical edges required only 2 levels of orientation whereas the diatom texture required at the least 18 levels (each $10°$ wide). It was felt that 20 or more orientation levels would be sufficient for most textures.

The number of spatial frequencies was changed by modifying the size of the discrete frequency transform. In all cases the accuracy of the result was not affected but different amounts of the textures were captured. For a few spatial frequencies only a small piece of the texture would appear in the mosaic, and for many frequencies a large area of the original would appear.

Does the above relate to the accuracy of the primary visual cortex ? Hubel and Wiesel[9] found the accuracy of the orientation neurons to be approximately $10°$. But there is a major difference between the FORP model orientation sensitivities and those of the primary visual cortex. The orientation sensitivities of neurons overlap much like the spatial frequency sensitivities. It can be shown that this may improve the actual orientation accuracy in the hypercolumn by quite a bit. In terms of frequency, nothing is lost as long as the local analysis region size changes with frequency accuracy. It has been shown that the receptive field sizes in the primary visual cortex are indeed proportional to their frequency bandwidth.[1]

---

† D'Astous 1983, p. 55

A model of the hypercolumn in the primary visual cortex that has many of hypercolumn traits has been presented. The model seems to be useful for characterizing texture but more texture samples need to be analysed.

Application of the Visual Cortex Model

Our visual system, if given the information contained in FORP space, could extract further information to characterize different textures. Neurons could be connected to these frequency and orientation selective neurons in such a way as to extract texture based features.

Michael[10] has described the method of inter-neuron information transfer. The output of a nerve cell is transmitted along its axon and is then connected to the input of another nerve cell via a chemical junction called a synapse. The synapses can either inhibit the nerve cell or excite it. The receiving neuron performs a type of summation or integration of all its synaptic input which results in an overall excitation level. If this excitation is above a threshold then the neuron will fire and transmit its excitation to other connected neurons. If the total inhibitory input is larger than the total excitatory input then the neuron will be inhibited from firing. Each synapse can also have an associated weight or a multiplication factor that emphasizes or de-emphasizes certain neural inputs.

This structure of inter-neuron connection is an effective means of constructing pattern recognizers. When a certain pattern of excitations is present on the axons of the input neurons, the processing neuron can recognize it and fire proportional to the strength of that pattern. In the context of the FORP cortical model, higher level neurons will be able to recognize certain patterns of firings of the frequency and orientation neurons, and as much as these patterns correspond to texture, these neurons will recognize textures.

Figure 7 depicts a simple realization of such a pattern recognizer. The hypercolumn is represented by a grid of small boxes. Each box corresponds to a neuron that is sensitive to a particular spatial frequency and orientation in one local area of the visual field. These neurons output to feature detector neurons through inhibitor and excitor synapses. The feature neurons shown can recognize very simple patterns in the FORP space, and may also receive either inhibitory or excitatory input from neighbouring hypercolumns. In Figure 7 the *Feature 1* neuron will detect a pattern of all one orientation with a strong low frequency component, a weak second frequency component, and a strong third frequency component. The *Feature 2* neuron will detect a pattern of one strong high frequency component and weak components on adjacent frequency and orientation neurons. These particular patterns might correspond to a particular texture or to a particular characteristic of textures.

Looking at Figures 3, 5, and 6 which are the FORP space representations of oriental rattan, diatom and J-Cloth™ respectively, it is easy to see how the neural mechanism described above could pick out patterns characteristic of each of these textures. Many more neural connections and some careful choice of inhibitory and excitatory synapse weights would be needed. A feature detector neuron could be designed to detect each of these textures since they are so different in FORP space. The implication is not that the primary visual cortex actually has a neuron for every type of texture, instead the suggestion is that a computer or electrical realization of such a mechanism could have separate texture recognizers if that was the end goal. In the visual cortex, the first level of feature extraction neurons would deal with texture attributes instead of specific textures. Second level neurons could combine these texture attributes to further specify texture. This hierarchical organization would be more general and more flexible than having textures recognized at the first levels.

One potential problem with this feature extraction mechanism is its sensitivity to orientation. A feature neuron that fires for a texture at one orientation may not fire when presented the same texture at a different orientation. People have little difficulty identifying the oriental rattan texture regardless of its rotation. Perhaps ours features should be orientation invariant.

Jolicoeur[11] demonstrates through perceptual experiment that rotational invariance need not exist in low level visual processing. He states, "... this experiment reveals a clear-cut effect of orientation on identification time. ... These results allow us to argue against a general model of pattern recognition based solely on the extraction of 'orientation-invariant features'."†

If texture is to be used in segmentation then orientation differences might be useful. Consider a cube with identically textured surfaces. When viewed, the main difference between the texture of the cube faces will be orientation, and hence orientation differences will play a major role in segmentation of the cube faces. Given these two insights, it would appear that these low level texture features need not be orientation invariant.

Segmentation

This feature extraction model can be extended to perform rudimentary texture based segmentation. Imagine the output of these feature cells being coded into a grey level depending on which cells were maximally excited. Then place these grey levels in a retinotopic map and generate a form of two dimensional image. Segmenting this image using grey level techniques such as thresholds,

---

† Jolicoeur 1985, p. 293

region growing, or edge detection is equivalent to segmenting the original visual scene by texture. The most probable candidate in human vision is segmentation via edge detection and shape recognition. Edge detection mechanisms that exist at this cortical level could be very similar to those found in the retina. An edge at this level will correspond to an edge between differently texture regions.

A combination of intensity edge maps, texture edge maps, colour differences, motion detection, and binocular depth cues can all be used in separating objects and performing image segmentation. Marr[12] discusses the use of multiple visual processes in providing accurate and stable decisions about surfaces. Each process involved in segmentation would provide its best information on surface boundaries, but only when all evidence is studied and judged for strength and weakness will the segmentation process decide on the final separation of constituent objects.

Summary

The image segmentation computer of the future will need to use various processes to mimic human success. It will use intensity and texture edges, colour differences as well as motion and depth cues. We have proposed a model of the primary visual cortex in which texture discrimination and segmentation can be performed. Demonstrations of the model indicate that similar textures are represented similarly and dissimilar textures dissimilarly. With extensions of further neural processing levels it may be possible to construct an effecient and effective texture segementation processor.

References

1. Pollen, D.A. and Ronner, S.F., Visual Cortical Neurons as Localized Spatial Frequency Filters, *IEEE Trans. on Systems Man and Cybernetics*, pp. 907-916 (September/October 1983).

2. Julesz, B., Textons, the elements of texture perception, and their interactions, *Nature* **290** p. 91 (March 1981).

3. Kimchi, R. and Palmer, S., Separability and Integrality of Global and Local Levels of Hierarchical Patterns, *Human Perception and Performance* **11**(6) pp. 673-688 (December 1985).

4. Berlucchi, G. and Sprague, J.M., The Cerebral Cortex in Visual Learning and Memory, and in Interhemispheric Transfer in the Cat, pp. 415-440 in *The Organization of the Cerebral Cortex*, ed. Schmitt, Worden, Adelman and Dennis, MIT Press (1981).

5. Hubel, D.H. and Wiesel, T.N., Uniformity of Monkey Striate Cortex: A Parallel Relationship between Field Size, Scatter and Magnification Factor, *Journal of Comparative Neurology* **158** pp. 295-306 (November/December 1974).

6. Brodatz, P., *Textures: A Photographic Album for Artists and Designers*, Reinhold, New York (1968).

7. Wolberg, L.R., *Micro-Art*, Harry N. Abrams Inc., New York (1970).

8. D'Astous, F.T., Textural Feature Extraction in the Spatial Frequency Domain, Ph.D. Dissertation, University of Waterloo (1983).

9. Hubel, D.H. and Wiesel, T.N., Sequence Regularity and Geometry of Orientation Coloums in the Monkey Striate Cortex, *Journal of Comparative Neurology* **158** pp. 267-293 (November/December 1974).

10. Michael, C.R., Retinal Processing of Visual Images, *Scientific American*, (May 1969).

11. Jolicoeur, P., The time to name disoriented natural objects, *Memory and Cognition* **13**(4) pp. 289-303 (July 1985).

12. Marr, D., *Vision*, W.H. Freeman and Company, San Fransisco (1982).
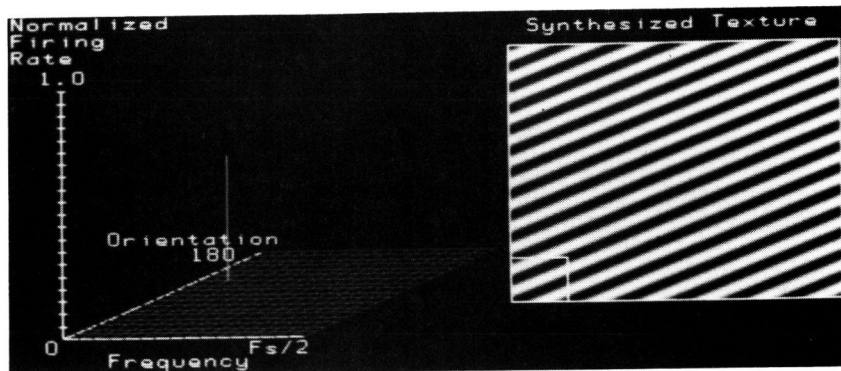
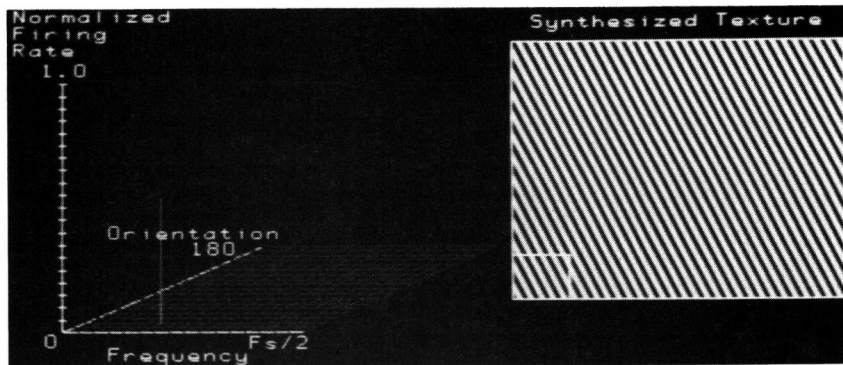Figure 1 : Single Point in FORP Space

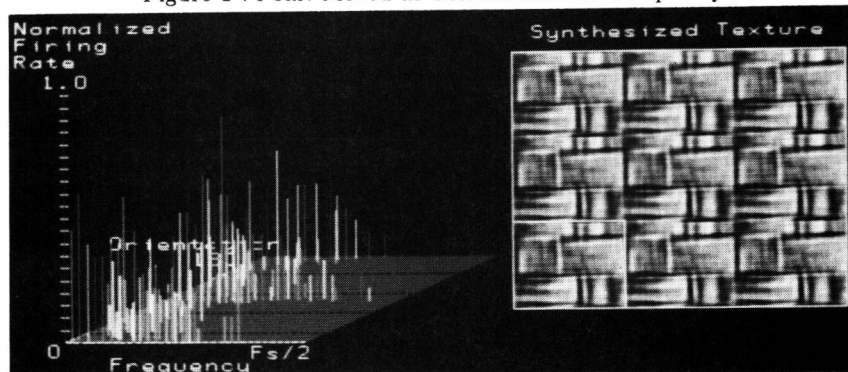Figure 2 : Point Moved in Orientation and Frequency

Figure 3 : Oriental Rattan - FORP Representation
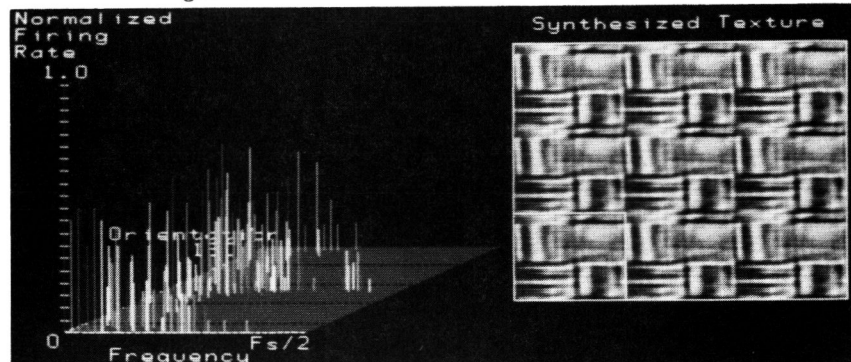

Figure 4 : Shifted Oriental Rattan - FORP Representation
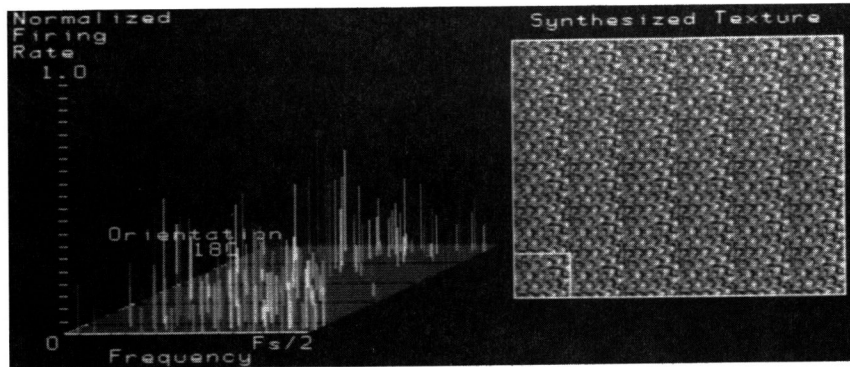
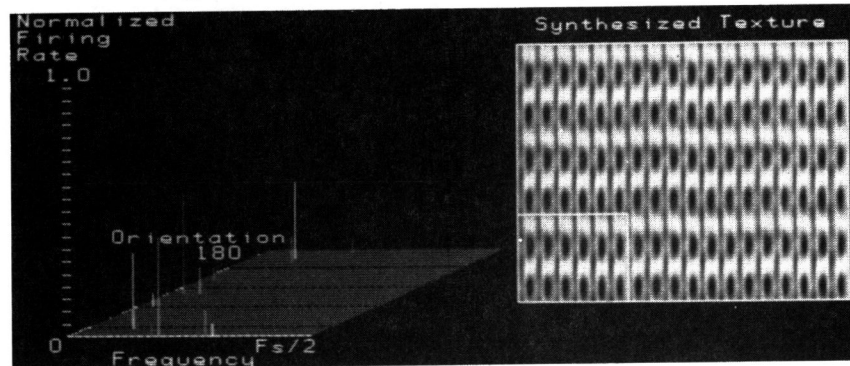Figure 5 : Diatom Texture - FORP Representation

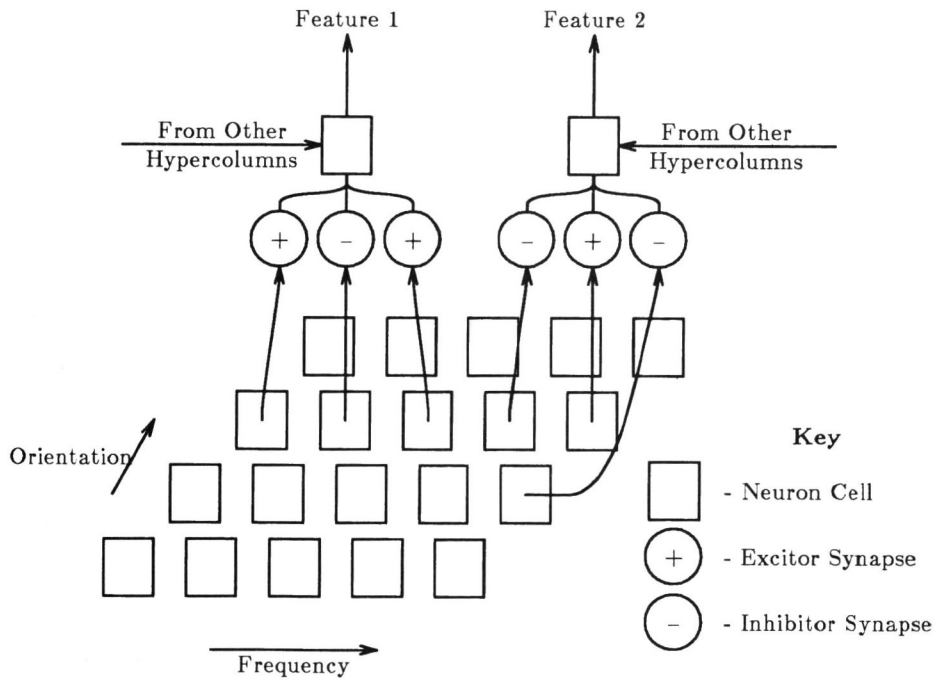Figure 6 : J-Cloth™ Texture - FORP Representation

Figure 7 : Simplified Model of Hypercolumn Feature Extraction