A Multiscale Description of Image Structure for Segmentation of Biomedical Images

James M. Coggins
Medical Image Display Research Group
Computer Science Department
University of North Carolina at Chapel Hill
and
Center of Excellence in Space Data and Information Science
NASA Goddard Space Flight Center, Code 630.5

Abstract

A new representation of image intensity structure, the Multiscale Orientation Field, is described and its application to biomedical image segmentation is discussed. The MOF is composed of orientation vectors at every pixel and at multiple scales. The vectors are computed from the outputs of an Artificial Visual System composed of spatial filters whose design is described. A new abstract feature space for orientation measurement is defined and a filter sensitivity function that yields a desirable mapping into that feature space is described.

1. Object Definition in Biomedical Images

An image is an uninterpreted recording of a spatial intensity distribution. In order to visualize or measure objects in the world, the image must be interpreted by labeling image regions according to the objects they portray. Object definition is a prerequisite to measurement, classification, recognition, or visualization of objects from image data. Object definition in biomedical images is particularly challenging because of accuracy requirements and the variety of shapes and scales of objects present.

Automatic labeling of image objects is beyond the capabilities of computer vision, and is likely to remain so. A more practical and appropriate objective for computer analysis of medical images is to generate a multiscale description of image structure that can be manipulated using tools such as the Image Hierarchy Editor [1]. The image structure description provides a language for human-computer interaction based on regions rather than on pixels or edges. Objects are defined interactively from their component regions. For such an interactive tool to be effective, we require a labeling of sensible regions and relationships among them.

Multiscale decompositions of images have been recognized as providing certain efficiencies and for allowing feature detection at multiple scales [2]. This paper argues that multiscale decomposition is a powerful operation in its own right, especially if the notion of "scale" is generalized to include measurement dimensions other than scale.

Filter set

Multiple Convolutions

Recombination algorithm

Figure 1: Artificial Visual System structure.

An image is processed by a linear multiscale filtering stage, followed by one of a small set of nonlinear recombination algorithms specific to the visual task and independent of the image data.

Section 3 introduces the pattern recognition approach we apply to pixels to infer image region structure. Section 4 presents a computing paradigm called Artificial Visual Systems that is used in Section 5 to derive a new representation of image structure, the Multiscale Orientation Field. Design of the filters for generating the MOF is described in Section 6, and methods for visualizing the MOF are described in Section 7.
2. What is different about this approach?

Efforts to represent image structure have usually attempted to abstract the image into a representation reduced in size but containing, in some sense, equivalent information. This approach, which leads to a "2-1/2D sketch" [3] or an "intrinsic image" [4] or other summary structure [5], attempts to concentrate the image information into relatively few pixels or into a few higher-level structures such as edge segments, ridges, blobs, textured regions, and the like. This approach quickly maps the image structure into symbolic objects on which AI tools can be applied [6]. Attempts to concentrate image information in a few symbols have failed to provide a convincing, generalizable solution to any image analysis problem, even on static, 2-D images such as photomicrographs of tissues or CT and MRI slices. Moreover, the approach has led to ad hoc engineering efforts instead of developing a sound scientific foundation on which future results can be built.

Rather than attempting to concentrate a representation of image structure in a few symbols, the Multiscale Orientation Field is a distributed representation in which redundant information is created in order to yield more robust inferences of region structure based on statistical pattern recognition operations (e.g. feature space mappings, clustering, and classification) applied at the pixel level. The regions obtained are suitable for interactive object definition. The use of statistical pattern recognition at the pixel level distinguishes this work from the pioneering work of Pizer and his students [7,8], who describe image structure using the Intensity Axis of Symmetry, a rich data structure that explicitly represents symmetry relationships in image regions.

3. Pattern Recognition for Image Analysis

The essence of pattern recognition is to map objects into points in an abstract feature space such that the structure of the feature space (the distance metric and spatial probability densities of object classes) reflects properties of the real objects [9]. The crucial challenge facing users of this paradigm is to define features that make sense both in relation to the objects and relative to each other. Typically, the latter consideration is ignored: incomensurable features arbitrarily defined yield feature spaces having no meaningful content or structure, and embodying arbitrary assumptions concerning the scale, distributions, and relative importance of the features and classes. This arbitrariness encountered in statistical pattern recognition studies conflicts with the desire for understandable, quantitative, reliable, robust inferences from the feature space and has led the computer vision field to seek other inference procedures, such as symbolic AI methods, that in fact face even more intractable difficulties.

The Artificial Visual System approach described below creates inherently comensurable, meaningful feature spaces. The Multiscale Orientation Field is a product of applying this approach to the problem of image structure measurement and representation.

4. Artificial Visual Systems

An Artificial Visual System [10,11,12] (Figure 1) consists of a series of filters that decomposes an image along a measurement dimension (spatial frequency, orientation, x- or y-direction, electromagnetic spectrum, time, brightness, etc.) in order to obtain at each pixel a vector of features, one from each filter. The feature vector maps each pixel into a feature space where a nonlinear recombination algorithm infers properties of the pixel or its neighborhood. The approach was inspired by filter-based models of the early human visual system [13,14] and has been reinforced by the image analysis techniques used by astronomers and geographers [14,15] to extract numerous properties of a scene from multispectral image sources. This section will motivate and develop the AVS paradigm.

Consider a single-valued stimulus moving through a measurement dimension. This notion can be made concrete by imagining the measurement dimension to be electromagnetic wavelength and the stimulus to be a spot of monochromatic light. The movement of the stimulus through the measurement dimension corresponds to the light smoothly changing color through the spectrum from red to blue.
Measurement of the stimulus value at some point in time might be accomplished by examining the output of a series of filters (Figure 2). A filter is a function defined on the measurement dimension that maps a stimulus or a pattern of stimuli into a number equal to

\[ F(s) = \int_{-\infty}^{\infty} f(\mu)s(\mu) \, d\mu \]

where \( \mu \) is the measurement dimension, \( s \) is the stimulus and \( s(\mu) \) is its \( \mu \)-spectrum, and \( f(\mu) \) is the filter sensitivity function. If one understands the structure of the set of filters and, thus, how the series of numbers is derived, one can, perhaps, recover the nature of the stimulus from the pattern of numbers output by the filters.

Figure 2: Measurement with a Set of Ideal Bandpass Filters

Figure 2 shows a sequence of ideal bandpass filters sensitive to adjacent bands of the measurement dimension. This filter set can determine the value of the stimulus to within half the width of the filters. To obtain more precision, we can build a set of filters with smaller pass bands at the expense of either building more filters or losing part of the range of the filter set. However, building more narrowly-tuned filters is usually more expensive than building wider filters.

Figure 3 shows a ramp filter sensitive to a particular range of stimuli just as our eyes are sensitive to a particular range of the electromagnetic spectrum we call "visible light". The filter has a different sensitivity for each possible stimulus value. The filter is drawn with a linear sensitivity function, but the linear shape is not essential; any strictly monotonic sensitivity function will provide a unique response to any stimulus value in the region of sensitivity.

The linear shape does simplify the computation of the stimulus value from the filter output. If we know that the stimuli all have the same intensity, we can compute the exact value of the stimulus by using the filter’s output value to interpolate between its high and low limiting values. In Figure 3a, for instance, if we denote the filter’s low and high limits as \( l \) and \( h \) and the filter’s output as \( F(S1) \) then the value of the stimulus is \( l + F(S1) \frac{h-l}{3} \).

Filters with more complex sensitivity functions may require more complex calibrations to infer the actual stimulus value.

Now suppose we allow the stimulus intensity to vary. Stimulus \( S1 \) (Figure 3a) yields an output value of 1/4. Stimulus \( S2 \) (Figure 3b) has an intensity of 1/3 and the filter sensitivity at the value of the stimulus is 3/4, so the filter’s output is \( 1/3 * 3/4 = 1/4 \). Unfortunately, the filter’s response to stimulus \( S2 \) is the same as its response to stimulus \( S1 \). We conclude that the stimuli \( S1 \) and \( S2 \) are indistinguishable to this one-filter Artificial Visual System. The problem is that the filter maps stimuli with different intensity/value pairs to the same output values.

Our approach can be saved by constructing a new filter, \( F2 \), having a sensitivity profile as shown in Figure 4. Filter \( F2 \) is just a reversed version of filter \( F1 \) used in Figure 3. We characterize a stimulus by treating the output values of filters \( F1 \) and \( F2 \) as the coordinates of the stimulus in a 2-D feature space. Each possible combination of stimulus value and intensity is mapped by the filters into a unique point in the feature space: \( S1 \) maps into \( (1/4, 3/4) \) and \( S2 \) maps into \( (1/4, 1/12) \).
The ability of a pair of filters to tease apart a stimulus, as in Figure 4, so that we can recombine the features to infer properties of the stimulus is the foundation of perception and the source of the computational power of Artificial Visual Systems. This kind of spectral decomposition is a powerful computational tool used by earth scientists for ground cover classification, by astronomers to determine properties of distant stars and galaxies, and by our visual systems to create representations of color and of spatial structure over multiple scales.

The slope of a filter's sensitivity function determines how much change is required in the value of the stimulus along the measurement dimension to create a "just-noticable difference" in the filter response. The tradeoffs in filter design are consistent with what might reasonably be expected: increased precision of measurement requires a filter with higher slope, therefore more filters are required to cover a desired range of values on the measurement dimension. To counteract noise, one needs a larger quantity of more finely-tuned filters until the slope becomes so high that most of the differences in filter response are due to noise. Note that using more, finely-tuned filters does not decrease the noise level in the signal; rather, fine tuning increases the difference in filter response for a given difference in stimulus value. This may result in a system more robust against noise and with higher precision.

Instead of using two filters to cover some range of the measurement dimension, we can use a bank of filters as indicated in Figure 5. A single-valued stimulus of arbitrary intensity appearing anywhere in the range of the filters yields a unique pattern of responses from the sequence of filters. The intensity and position of the stimuli in the sensitive range can be deduced from the pattern of responses.

Since the filters in the AVS are overlapping on the measurement dimension, their responses are not independent. This redundancy can be exploited in the feature space mapping to make the resulting inferences robust in the presence of noise and distortion. Also, since the filters decompose the image in a uniform manner on the same measurement dimension, the resulting features are inherently comensurable, resulting in an intuitively understandable and mathematically tractable feature space. In fact, such a feature space was used by this author in an earlier project to correct for an image distortion by geometric manipulation of the feature space [12]. Such a facile use of the feature space is possible only if the feature space is intuitively and mathematically understandable.

Studies of the properties of artificial visual systems such as just-noticable differences are analogous to psychophysical studies of human vision with the advantage for investigators that objective comparisons of AVS feature spaces are possible since we have direct access to the internal representation of the stimuli. These representations are amenable to statistical analysis and evaluation using methods that are more meaningful than simply comparing classification results [16,17,18]. Such methods include the Hotelling Trace statistic and ROC analysis.
5. Multiscale Orientation Fields

Before defining the multiscale orientation field, we must explain the notion of an "orientation spectrum" illustrated by the three examples in Figure 6. This orientation dimension can be characterized as

```
+------------------+
|                  |
|                  |
|                  |
|                  |
+------------------+
```

Figure 6: Simple Images and Orientation Spectra

"undirected orientation about the center of the image". A line segment centered at the image center is a single-valued stimulus in this measurement dimension. A darkened pixel at an orientation \( \theta \) from the center of the image contributes to the spectrum at orientation \( \theta \). Thus, a longer line segment produces a larger spike in the orientation spectrum.

```
+------------------+
|                  |
|                  |
|                  |
|                  |
+------------------+
```

Figure 7: More Complex Orientation Spectra

Moving the stimulus through the measurement dimension corresponds to rotating the line segment smoothly about its center. The measurement dimension is circular and 180 degrees long since an undirected line segment rotated 180 degrees is indistinguishable from the original. This two-fold symmetry will be reflected in the interpretation procedure below.

This orientation dimension can be used to infer properties of image structures other than lines, as illustrated in Figure 7. Not only does the maximum of the orientation spectrum indicate the orientation of the principal axis of the oval, but the integral of the spectrum indicates its area, and the peakedness of the orientation spectrum is a measure of circularity (or eccentricity). Also, the shape of the spectrum can be used to distinguish ovals from other shapes, such as rectangles.

```
+------------------+
|                  |
|                  |
|                  |
|                  |
+------------------+
```

Figure 8: Mapping of the filters to the orientation feature space (a) spatial directions of the orientation filters (b) vectors in the orientation feature space corresponding to each filter

Rather than computing and analyzing the entire spectrum of a stimulus, we can apply an AVS as in Figure 5a to obtain a small set of redundant measurements of the orientation spectrum. The measurements map the stimulus into a polar feature space as illustrated in Figure 8 as follows: each filter output determines the length of a vector in the feature space oriented along an axis corresponding to the filter's center orientation (Figure 8b). Since the filters are identical except for their orientation shift, the axes of the feature space are evenly distributed about a circle. Due to the twofold symmetry of the orientation dimension and thus of the filters, the axes in the feature space are separated by twice the angle between the center orientations of the filters. This configuration allows the filter outputs to contribute both positive and negative information to the inference of principal orientation: the lack of a strong response in the 90 degree filter may constitute supporting evidence that the principal orientation lies in another direction.
The vectors in the feature space are summed to yield a resultant vector as shown in Figure 9. The orientation of the resultant indicates the principal orientation of the oval. The length of the resultant indicates the strength of the orientation preference in the resultant direction, which is a measure of eccentricity.

![Figure 9: Computing the principal orientation and eccentricity from the filter outputs](image)

This discussion has described the computation of the principal orientation about the center of an image only. Computing principal orientation vectors at every pixel gives an orientation vector field. To create a Multiscale Orientation Field, we compute the principal orientation vector at every pixel of an image at multiple scales.

Our studies have used six orientation kernels (Figure 5a) oriented 30 degrees apart with sensitivity functions that fall to zero 30 degrees away from the center orientation, at the center orientation of the neighboring filter. The orientation kernels are multiplied by a circular Gaussian kernel (Figure 5b) to define the scale of the filter set. The variance of the kth Gaussian is $\sigma_k = 2^k$ in pixel units, and $k=0...K$, where K is selected based on the image sample size and the size of structures of interest in the image. Typical values of K are from 6 to 10. Each of the 6K kernels convolved with the image is bandlimited in both orientation and scale.

The Multiscale Vector Field describes for each pixel the principal orientation and eccentricity in a series of circular (Gaussian) neighborhoods of increasing size. Comparisons of orientation vectors obtained at neighboring pixels and at neighboring scales can reveal orientation and scale coherence that is crucial for defining sensible regions in medical images.

### 6. Design of Orientation Filters

The linear sensitivity functions illustrated in Figure 5a have certain drawbacks that can be corrected with a more sophisticated filter design. To introduce this design, we return to the line segment measurement example.

Consider again a line segment smoothly rotating about the center of the image. We map the moving line segment to the circular orientation feature space using filters with linear orientation sensitivity functions. The locus in the feature space of the rotating line segment is a polygonal path as shown in Figure 10. The size of the polygon changes with the length of the line segment: longer line segments produce larger polygons.

![Figure 10: Locus of the mapping of a constant intensity stimulus to the orientation feature space](image)

Using linear sensitivity functions, the rotating vector maps into chords of a circle, so the length of the resultant vector and the derivative of the vector orientation with respect to line segment orientation both vary. Therefore, there is ambiguity in the interpretation of a given vector and some measurement error arising from that ambiguity. More precise measurement is possible if a constant-intensity, single-valued stimulus maps into a circle in the feature space.

Such a mapping is obtained if the sensitivity function of the orientation filters has the form

$$F(\theta) = \frac{\sin(\Omega - \Omega_0 \theta)}{\sin \Omega}, \text{ for } 0 < \theta < 30^\circ$$

$$= 0 \text{ otherwise}$$

where $\theta$ is the angle from the filter center orientation, $\omega$ is the angle between filter centers in the image space, and $\Omega$ is the angle between the axis vectors in the circular feature space [19]. Note that $\Omega/\omega$ is the number of
periods of the orientation dimension in a full circle, which equals the number of lobes in an orientation filter; there are $2\pi/Q$ orientation filters required to span the orientation dimension; and the orientation dimension has a period of $2\pi/Q$.

As $Q$ decreases, a smaller portion of the sine wave is used to create the filter, making the sensitivity functions approach linear shapes. Decreasing $Q$ implies that more filters are used to span the orientation dimension.

To create an orientation filter, express the sample point in polar coordinates from the origin of the image, compute the angular difference between the sample and the filter center orientation being careful to express this difference modulo the length of the orientation dimension, and compute $F(\theta)$ for the sample point.

7. Visualizing the MOF

To evaluate the MOF, we have developed several visualization tools. We can manually select certain families of pixels and then superimpose plots for all pixels in a family showing the responses of isotropic Gaussians across scale, the magnitude of MOF vectors across scale, and the orientations of the MOF vectors across scale. Analysis of medical images using this tool has revealed that there exist pixel families with similar profiles in the isotropic Gaussians but different profiles in the MOF. Such pixel families have similar gray levels but are contained in regions having different orientation structures. Also, there exist different pixel families that have similar magnitude profiles in the MOF but quite different profiles from isotropic Gaussians and from the orientation of vectors in the MOF. This suggests that the orientation information in the MOF is valuable for region labeling. These results also suggest that combining information from isotropic Gaussian filters with the MOF may yield better labelings than with either alone.

We can also display the orientations of MOF vectors at a particular scale as an image. We have found in studies with both artificial and biomedical images that the orientations in a region of uniform gray level are random until the scale increases enough for the filter kernels centered over a pixel to interact with the region boundary. At that scale, the orientations of the MOF vectors snap into alignment with the orientation of pixels nearer the boundary. As the scale increases, more pixels near the center of the region join this alignment. The locations where opposite boundaries first interact is the symmetric axis of the region.

Having visually identified different patterns of responses in different families of pixels, we are developing procedures to characterize those families using the responses at each pixel to isotropic Gaussian filters and in the MOF as features for further analysis using classical statistical pattern recognition operations.

8. Conclusion

Other researchers are investigating oriented filters for computing optic flow [20,21] and as features for general spatial vision [22]. Hsieh [23] uses oriented Gaussian filters in a connectionist image segmentation scheme that yields ordinal inferences of orientation, not quantitative inferences as in the MOF.

We have presented a new approach for representing multiscale image structure based on Artificial Visual Systems that decompose an image along the dimensions of orientation and scale. The Multiscale Orientation Field represents the orientation and eccentricity of image intensity in multiscale neighborhoods about each pixel. A feature space and a filter design for producing MOFs were described.

We have presented an approach to representing image structure using statistical pattern recognition techniques at the pixel level. The representation of structure is distributed throughout space and scale, creating and exploiting redundancy to achieve robust inferences of structure.

Acknowledgements

This research was supported in part by the NASA Center of Excellence in Space Data and Information Science through the Universities Space Research Association.
References


