When we know enough to describe brain processes capable of giving perception in machines, the problem of consciousness may be solved - or shelved, as the substance of matter is ignored in physics.

\- *Richard Gregory* - Eye and Brain

Chapter 1

Introduction

Computer vision often starts with multiscale representations of images that attempt to capture image structure. The multiscale differential geometric operators used in this type of approach measure local image properties such as contrast and curvature. When such operators are applied across a varying range of apertures or local neighborhood sizes, a deep understanding of image structure can be achieved.

Multiscale models for computer vision are motivated by our knowledge of the early processes of attention, focus, and object definition in mammals. Laboratory analysis has revealed a variety of visual receptors that detect and measure boundaries, orientation, linear motion, contrast, and relative size. These receptors are spread throughout the visual field in varying concentrations. The receptive fields of these visual operations vary significantly in the amount of the visual field that they cover. Also, the various fields are known to overlap one another. These receptors provide the later stages of the visual system with a rich assortment of inputs from which to organize and formulate conclusions.

Multivariate images present a problem for multiscale computer vision. In some cases, the separate values within a pixel are incommensurable; that is, they lack a common basis for measurement. Where such conditions exist, the multiple values at an image pixel cannot be considered a vector value and the language of differential geometry is no longer applicable.

Another common approach to computer vision uses statistical pattern recognition. This approach begins by performing a statistical analysis of values measured from an image. Data gathered from the image are analyzed, and statistical trends, correlations, and probabilities are extracted. The resulting statistical models of the image are then used to help frame decisions in the later stages of processing. Such statistical approaches provide powerful means for making decisions about image structure in the presence of noise and in the cases of multivariate data.

This dissertation presents new models for early computer vision. It synthesizes methods from multiscale differential image geometry and statistical pattern recognition to form a new class of multiscale statistical operations on local image intensities. The goal is to provide new mechanisms for the analysis of multivalued images. Presentation of the problem addressed by this research requires some background on the organization of computer models of vision and the statement of some basic assumptions about the structure of computer vision systems.

1.1. A multiscale approach to computer vision

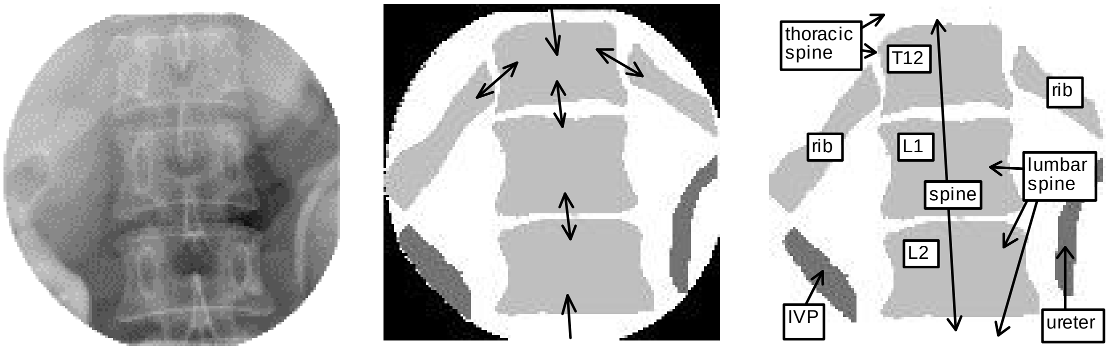
A *feature* is an attribute of an image, perhaps measured at a particular location. Feature measurement is a fundamental first step in any image processing task. Typical features or attributes of an image include intensity, the gradient direction, the gradient magnitude, color, texture, the signal-to-noise ratio, boundariness and medialness [Morse 1993], and other measurements that may be computed at a point in image space.

In the presence of noise, however, feature measurement requires regularization. This is often achieved through area averaging or smoothing of the measured values, reducing the noisiness while sacrificing resolution. Such smoothing techniques add a dimension of scale to any measurement operation. The amount of smoothing (or the aperture of the scale operation) used to attain a desirable result is a subject being actively explored today.

Areas of similar feature value may be aggregated into regions or segments. A *region* exhibits feature continuity if measurements of features (intensity, texture, color, gradient, etc.) within its boundaries are changing smoothly. In many models of images, sharp discontinuities or step edges in some feature value are often expected at region boundaries. Analytical descriptions of regions often include the measurement of functional features such as the medialness or boundariness of pixels within a region.

Often it is difficult to identify and separate boundary conditions from the effects of noise within the image. Noise removal often softens boundaries, creating problems in the identification of the borders of regions. On the other hand, boundary measurements that attempt to amplify region-delimiting discontinuities often exacerbate noisy conditions. Thus, the smoothing or scale operator is often applied across a range of values, and the resulting multiscale representation is studied.

I include the recognition and representation of the relationships among regions (connectivity, feature similarity, etc.) in the process of *segmentation*. For example, consider a digital radiograph of a portion of a human torso (Figure 1.1). It is possible to capture individual vertebrae during segmentation. A higher order segmentation stage links the bones into their skeletal structure based upon their connectivity. During the classification stage, components are labeled according to a variety of semantics. Pixels having certain attenuation properties are identified collectively as “bone.” Individual bone segments may be classified by their clinical names (metatarsal, iliac, femur, L1, L2, L3, etc.). Finally, the higher order distinction may be made of the connected segments into their hierarchical structure; the twelve thoracic vertebrae (T1, T2, ..., T12) denote the *thoracic region* of the spinal column, the five lumbar vertebrae (L1, ..., L5) denote the *lumbar region* of the spinal column. There is even a higher level abstraction joining these separate regions; the combined cervical, thoracic, and lumbar regions along with the sacrum and coccyx comprise the “spine.” Some regions of high attenuation are rejected for semantic reasons. In this case, the kidney and the ureter have artificially elevated attenuation coefficients because of the application of a contrasting agent.



a. original image b. segment mask c. classified segment mask

Figure 1.1. A segmentation example. (a) the original digital radiograph, (b) an image mask denoting segments, and (c) the classified segment mask, showing the hierarchical semantic organization of the skeletal system.

The separation of the image into partitions or regions is difficult in the example above. Variations in shading, opacity, curvature, and other forms of interference make difficult successful automatic segmentation. Even the indivisible pixel measurement itself may reflect multiple characteristics, sharing many partial attributes within the area or volume encompassed by the pixel sample. This problem of *partial-voluming* is a result of discrete sampling. Region delimiting methods involving binary decisions are therefore often ill-posed, giving rise to inaccurate analysis and unacceptable artifacts.

Partial-voluming requires that fuzzy truths about image features need to be represented and measured. Thus, statistics often play a significant role in image analysis, with errors and probabilities measured of image features. The effects of partial-voluming combined with the frequent need to analyze multivalued data compels the use of multiscale statistical techniques. Statistical analysis necessitates accurate, robust, and versatile feature measurement that facilitates error analysis and probability computation.

1.2. An integrated approach to early vision

From the perspective of object definition, feature measurement should capture information about the local statistical structure of the image in preparation for assessments of feature continuity. This incorporation of statistical methods directly into feature measurement is in contrast to the more common two step approach of first measuring features across an image and then applying a statistical analysis to determine the distribution of the observed values.

Beyond the capturing of the probability distribution of the image intensities, multiscale statistical measurement also can be shown to capture local spatial trends of feature values. Image geometry can therefore be studied through statistical measurements. Stated informally, rather than apply a statistical analysis to features designed to measure image geometry (i.e., the statistics of geometry), I suggest using a multiscale statistical analysis to directly measure the geometry of images. That is, I propose to study *statistics AS geometry*.

1.3. Driving issues

This dissertation describes a new approach for feature measurement in image processing; it replaces or supplements existing building blocks in the foundation of computer vision. It is a synthesis of applied statistics and computer science to address the problems of preprocessing two-dimensional digital images to allow improved object definition. The result is a new form of multiscale statistical feature measurement. These features are developed through a combination of probability theory, traditional statistical pattern recognition, and multiscale geometry. The following discussion lists a progression of the issues that arise from region identification and noise reduction. These issues are the framework for this research.

Scale – The term ‘feature’ has been defined as an attribute of a sample of an image. In the simplest terms an individual pixel is the smallest atomic sample. However, a pixel is only a single observation from a broader population. If the image contains noise or representation error, it may be difficult to reliably measure features and their trends across an image from collections of individual sample measurements. Indeed, given a continuous probability density function to describe the spread of the noise in an image, the likelihood of encountering a particular feature value is zero. Decisions based upon a single observation or pixel value are inherently unstable. Some measurement aperture or *scale* must be applied when analyzing noisy digital images, thus including a local population of pixels in any feature measurement.

Moments – In the context of aggregations of pixels a powerful metaphor for describing images and image samples based on probability theory may be employed. Using scaled measurements, it becomes possible to discuss a pixel and its neighborhood in terms of its expected values. Collections of pixels from the image population may be evaluated, and statistics (a body of descriptive measurements about the collection of pixels) can be generated about the pixel group. A series of statistics of an image region that can be estimated from a weighted collection of image pixels are its central moments of intensity. By ‘moment,’ I mean the statistics of the distribution of pixel intensities. That is, I refer to the classical definition of the term ‘moment’ as used by practitioners of probability theory and not the moments of location in an image found recently in the image processing literature.

The sample mean and consequent central moments are useful statistics in describing the distribution of a pixel population. The shape of the distribution of pixel values from a neighborhood about an image location is indicative of many important features of the image including the likelihood of a nearby boundary, local continuity of features, and the approximate variance in the signal. Additionally, I contend that moments themselves are vital features that may be exploited in image understanding.

Multiscale Analysis – This thesis assumes that regions are spatially contiguous. That is, individual pixel values are spatially correlated. Under this assumption, and given the stability of scale space, it is advantageous to employ a spatial sample of each pixel and its surrounding region to determine the local image statistics. Since the actual structure of the image is not known *a priori*, a weighted spatial sampling kernel should be applied at several different measurement apertures. Multiscale statistics reflecting the local probability distribution of image intensity are computed at each pixel location over a range of varying scales. The continuous space of image statistics at all scales can be combined into a scale-space analysis. I generalize local central moments for scale-space analysis and project their applications.

Uncertainty – The process of *partial voluming* does not allow a simple binary classification and is one of the fundamental issues addressed in this dissertation. All stages of the visual process must handle this fuzzy property of pixels containing a variety of feature expressions from early filtering, through segmentation, and through classification. I address this issue by invoking measurements and approximations of the distribution of pixel intensity values to compute the probable inclusion of each pixel within different possible regions.

Geometry – Detecting and measuring spatial variations of feature values are essential to the combined process of segmentation and classification. Feature continuity, spatial contiguity of object segments, and the varying sampling aperture of image statistics generate the necessity of understanding the multiscale geometry of a given image. Identifying the traits of feature continuity and the discontinuities that define segment borders is a problem of geometry. Defined broadly, *geometry* is the study of properties of given elements that remain invariant under specified transformations. By introducing directional components into multiscale moments of intensity, I achieve geometric measurements of image structure. I study the invariances of directional statistical moments and incorporate them in image analysis.

1.4. Thesis

This research synthesizes the ideas of scale space and statistical pattern recognition into a coherent framework for noise reduction and segmentation. The algorithms that are generated from this approach are demonstrated to be robust with respect to variations of contrast and brightness within the image as well as invariance with respect to rotation, translation, and changes in scale.

Thesis:

> Local image behavior can be usefully measured by first imposing a Gaussian neighborhood operation on the image and then estimating central moments of the local probability distribution function from the resulting weighted samples of image intensity.

The resulting statistics represent a novel form of multiscale analysis. Measurements of the statistical behavior of images can now be made at different scales.

The primary goal of the development of multiscale image statistics is to enable automatic segmentation by nonlinear diffusion. The intent is to analyze images, extracting the necessary features to set the parameters of a nonlinear diffusion equation. Multiscale image statistics are also expected to improve other segmentation methods.

1.5. Overview

As a work of synthesis, this dissertation approaches the problem of feature measurement from many separate domains. It is presented as a progression of topics. After establishing some background on scale-space theory and some essential probability identities, I explore the statistics of multiscale derivatives. Next, I develop a series of local image statistics, specifically, local central moments, that describe the behavior of image intensities within a local neighborhood. Subsequently, I show that local central moments can be calculated either over isotropic neighborhoods or with a directional component. Directional elements enable central moments of image intensity to describe complex spatial behavior within the image. Both the isotropic and directional families of multiscale intensity moments are generalized to describe multiparameter images, images with more than one value per pixel. Finally, a chapter describing future work in this research area speculates on how local moments can be used in nonlinear diffusion systems to preprocess images for further analysis by statistical or structural pattern recognition algorithms.

Chapter Two offers the pertinent background in image processing, scale space, probability, and statistical pattern recognition that serves as the foundation of this research. This chapter emphasizes those elements that are crucial to and supportive of the thesis and introduces a vocabulary and notation that will be used throughout this writing. Chapter Two is divided into multiple sections to encourage the reader to skip familiar material.

Chapter Three is a statistical analysis of multiscale derivatives of noisy images. It discusses the propagation of spatially uncorrelated noise through scale space and the effects of noise on the family of scale-space differential invariants.

Chapter Four introduces local central moments of intensity based on an isotropic Gaussian neighborhood function. These moments represent a family of operators that measure the local behavior of the probability distribution of image intensities. The construction of this type of image decomposition is justified by modeling images as stochastic processes. These statistics are generalized from scalar-valued to multivalued images. The chapter also explores the propagation of uncorrelated noise from the original image to the measured local central moments.

Directional elements of multiscale image statistics are described in Chapter Five. The earlier isotropic moments of Chapter Four are extended to show spatial and orientation biases in image structure. Methods for two-valued 2D image analysis through multiscale directional central moments of image intensity are suggested through the application of canonical analysis of the covariances of the separate image values.

Chapter Six explores the ramifications of these ideas and their extensions to related problems in computer vision. I suggest extensions of multiscale statistics, including analysis of the propagation of uncorrelated noise through nonlinear scale space, the adaptation of Pearson’s taxonomy of probability distributions to a scale-space framework, and the modification of the Kolmogorov-Smirnov test to multiscale image statistics. I also speculate on several applications of multiscale image statistics as aids in segmentation. I present examples of multiscale image statistics in boundary estimation and segmentation by anisotropy invariants. I also revisit some earlier work in statistically based variable conductance diffusion (VCD) and supply some conjectures on how multiscale statistics might be incorporated into the scalar-valued and multivalued zeroth order VCD equation. A summary and discussion presented in this chapter ends the dissertation with a look to the challenges that remain.

1.6. Contributions

This dissertation includes three substantive original contributions. The first development is original work on the propagation of uncorrelated noise through linear scale space. Using the scale-space normalization suggested by Eberly, results show that for all given levels of initial intensity of noise, the absolute error in the multiscale derivative decreases between zeroth and first order measurements. The level of propagated noise increases thereafter with increasing order of differentiation; however, it remains less than the initial error until derivatives of the third or fourth order are taken. This finding brings into some question the common wisdom that low order differentiation badly propagates noise.

The second original development is multiscale image analysis based on local statistics rather than on geometric measurements such as local derivatives. These isotropic multiscale image statistics are significant in several ways. First, image analysis based on multiscale image statistics can easily be made invariant to linear transformations of intensity as well as to spatial rotation and translation. This trait makes the identification of an object independent of the absolute brightness of the object as well as independent of its orientation and position within the image. Also, multiscale statistics provide a means of analyzing multivalued data where the data channels within the image are incommensurable (i.e., they have no common metric for measurement).

The third significant contribution presented in this dissertation is the development of directional multiscale image statistics. These central moments reduce bias in noise characterization that arises from the gradient of the image function. Singular value decomposition of directional covariances produces principal axes indicating maximum and minimum spread of the directional probability distributions. These orthogonal directions and their corresponding eigenvalues can be used to normalize measurements made of local intensities. Multivalued image analysis is also possible through directional covariances; normalized multilocal coordinate systems based on covariances can enable comparisons where the lack of common distance metrics makes vector analysis impossible.

Beyond these findings, I foreshadow a well posed means of computing parameters for traditional and non-traditional nonlinear diffusion equations directly from multiscale statistics of the local image intensity.