# Chapter 6

# Summary

The main goal of this dissertation is to provide the foundation for image analysis via geometric methods, in particular using the concept of ridges. I provide here a summary of the chapters of my thesis together with what I believe are my contributions to the theory.

## 6.1 Ridges

I have provided various formal definitions for ridges and discussed each definition from a geometric point of view. The definitions are discussed in terms of their invariance properties and locality of construction.

Ridges constructed according to the reverse gravity watershed definition are invariant with respect to spatial rotations, spatial translations, uniform spatial magnifications, and monotonic intensity transformations, but the construction is a nonlocal process. The nonlocality makes the definition suspect in applications where medial structures are to be found. What a viewer identifies as a ridge should not be affected by image structure at locations far away from the region of attention.

Ridges constructed according to the height definition are invariant with respect to spatial rotations, spatial translations, and uniform spatial magnifications, but are not invariant with respect to monotonic intensity changes. The construction is a local process. The lack of invariance under monotonic intensity transformations requires care in constructing ridges that are visually meaningful. Usually the true image data is monotonically transformed for display on monitors with a limited number of gray values. What a viewer identifies as a ridge might just be an artifact introduced by the transformation. Conversely, a ridge in the true data might be eliminated by the display transformation.

Ridges on the graph of intensity constructed according to the principal direction definition are invariant with respect to spatial rotations, spatial translations, but not with respect to uniform spatial magnifications and monotonic intensity transformations. The construction is a local process. For surfaces obtained as level surfaces of functions, the ridges constructed by this definition are invariant with respect to rotations, translations, and uniform magnifications of the space in which the surface is embedded. As in the height definition, some care is required when constructing ridges because of the lack of invariance under monotonic intensity transformations (for graphs). The principal direction definition appears to be most useful for analyzing $n$–dimensional surfaces embedded in $(n + 1)$–dimensional space where the surface information is purely spatial and not relating to image intensities. A contribution I have made for this definition is to show that the ridge construction can be performed without having to parameterize the surface of interest.

Ridges constructed according to the level definition are invariant with respect to spatial rotations, spatial translations, uniform spatial magnifications, and monotonic intensity transformations. The construction is also local. The definition is a special case of the principal direction definition whereby level ridges are the union of all principal direction ridges of the level surfaces of the intensity function. Since the level definition has all the desired invariances, it is a theoretically good choice for applications. However, in practice I have found that the height definition appears to produce more meaningful ridges than the level definition. A reason for this difference might be that the level definition requires computing higher order derivatives than does the height definition. Many of the supposed ridge points occur because of noise in the signal.

I have not fully investigated the implications of using the nonmetric definition for constructing ridges. This definition has the same invariance properties as the principal direction definition, and it is a local process. Computationally it is an attractive alternative to the principal direction definition, but I need to investigate whether or not the nonmetric ridges are visually meaningful and useful in applications.

Most of the ridge constructions found in the literature are applied to 2–dimensional images, the resulting ridges being 1–dimensional structures. A contribution I have made to ridge analysis is to provide ridge definitions for general $n$–dimensional images. Moreover, I have extended the notion of a 1–dimensional ridge structure to that of a $d$–dimensional structure, allowing image analysts a more general and flexible tool for studying images.

## 6.2 Object Construction

Ridges appear to play a major role in local object definition and recognition, but obtaining more global information about objects necessarily requires the concept of scale. My ridge flow segmentation algorithm is an attempt at providing such global information.

Segmentation based on reverse gravity watershed methods is not adequate. The primitive regions are built by constructing ridges via a nonlocal process. The identification of a region as "flanks" associated with a ridge should also be a local process, at least near the already identified ridges. Originally it was thought that the flows one should follow in identifying pixels with ridges are just the gradients of intensity. The problem with these flows is that for smooth intensity functions, the flows originate/terminate at critical points for the function (local extrema and saddles), but they generally do not intersect ridges transversely. Our intuition is that flows to a ridge should terminate orthogonally at a ridge, at least for the purposes of identifying pixels with ridges. In all of my local ridge definitions, I have indicated which flows should be followed, such flows always terminating orthogonally at a ridge. The flows are naturally related to the equations which define the ridges.

My algorithm for constructing image hierarchies is based on annihilations and merging of primitive regions through scale. As such, the algorithm captures global information about the objects in the image. Merging of primitive regions via nonlinear blurring signifies that previously distinct portions of an image at one scale are indistinguishable at a larger scale. The hierarchy construction therefore relates regions that are initially geometrically distinct, but records the scales at which they become indistinct. The idea of geometric similarity/dissimilarity in the algorithm may be useful in other multiscale segmentation schemes.

One implementation issue of interest is the segmentation of ridge structures into distinct components. I developed a thinning algorithm which is designed to take candidate

1–dimensional ridge structures and reduce them to a form in which pixel–thin curvilinear segments could be labeled. The algorithm works for any dimension image, preserves the topological properties of the input set (*i.e.*, it preserves the number of holes in the input set), and generally preserves the shape of the original set by using operations of mathematical morphology to thin the input from the outside–to–inside. The connected component routine I use also works for an image of any dimension and is quite efficient. These routines can be useful for other image analysis applications.

The segmentation algorithms I developed based on ridges and multiscale methods have proved to be successful in terms of an interactive, user–assisted computer environment. The current Magic Crayon prototype, which is an interactive tool for visualizing the hierarchy, will eventually become a tool used by clinicians.

## 6.3 Scale Space

I have shown how the basic assumptions about front–end vision systems lead naturally to a definition of scale space which requires specification of a metric. In most cases, the metric is data dependent and imposes a geometry on scale space which is non–Euclidean. Moreover, once the metric is chosen, the anisotropic diffusion process used to create multiscale image data is automatically determined.

For the scale space metric where conductivity is just the scale parameter, I have used the standard tools from tensor calculus to derive formulas that are necessary for making scale space measurements. These measurements exhibit invariance under spatial rotations, spatial translations, and uniform magnifications in space and scale since the arc length form has the same invariances. In particular, I have derived the formulas for computing distance between points in scale space, the gradient and the Hessian in the covariant sense, volumes of scale space regions, and principal curvatures and directions of surfaces.

I have also generalized the ridge definitions for Euclidean space to ones for scale space. The height definition for ridges remains essentially the same except that the Hessian matrix used is the one derived for functions defined on scale space. In Euclidean space, ridges are defined as those points for which the intensity function has local maxima when restricted to a linear (flat) subspace of the domain. In this restricted sense, the second derivative test is a test for definiteness of the Hessian restricted to the flat subspace. In scale space, ridges are defined in the same way, except now the "flat" subspaces are geodesic surfaces (surfaces of zero curvature). The principal direction definition also remains essentially the same except now the matrices representing the first and second fundamental forms are the ones which take into account the curvature of scale space. Similar to the Euclidean case, the scale space level definition for ridges is equivalent to applying the principal direction definition to all the level surfaces and then taking the union of the results. The invariance properties for the scale space ridge definitions are the same as those for the Euclidean space definitions.

The necessity for scale space ridge definitions is apparent in the application of finding cores of objects in images. Unlike the Euclidean case where ridges are defined on graphs of intensity functions, the important functions to consider in scale space are medialness functions. A core of medialness is naturally modeled as a ridge of the medialness function. Cores of objects should be invariant with respect to rotation, translation, and zooming of the objects. Scale space ridges satisfy these invariance properties because of the choice of metric. The consequences of imposing the correct metrics on scale space are far–reaching, and we expect the working out of their influence on image analysis applications to occupy us for years.

## 6.4 Future Research

For future research, I plan on investigating the following topics. Almost all of our applications require computation of the cores of objects. Although the geometric ideas I developed for scale space, namely ridge and valley structures, lead directly to construction of cores, there are a few issues of implementation to be dealt with. The sampling of scale as a geometric sequence creates a few problems in discrete derivative calculations.

Another issue of importance is labeling subcomponents of cores which are built as $d$–dimensional structures with $d > 1$. Such labeling is necessary to allow the construction of multiscale image hierarchies. Labeling subcomponents is not a trivial issue since the topology of a medial surface can be fairly complicated. Further research is needed on the topologies and singularities of cores. In particular, research is needed for understanding how primitive constructs such as curves and surfaces combine to form a complicated core which contains many holes and branching structures.

Using a single scale parameter, we can construct the core for a figure as a 2–dimensional structure embedded in a 4–dimensional space. The core contains information about figure shape and detail, and as indicated may be topologically complicated. I propose to extend my scale space analysis to handle a "recursive descent" description of shape. The core itself inherits its geometric nature from scale space and could possibly be represented compactly with yet another core, which conceivably has a simpler positional representation, but with more than one scale attribute attached to it.

One of our applications is object–based interpolation. The conventional method for interpolation of two adjacent slices in a 3–dimensional data set is to pair up pixels with the same $x$ and $y$ locations and interpolate their image intensities in the $z$ direction. This pairing is ignorant of the type of the pixels. It is meaningless to pair up a pixel representing heart tissue with one representing lung tissue. Pixels should instead be paired based on what type they are and/or what objects they come from. Initial studies have convinced us that the pairing must be based on spatial and scale information. However, the question of assigning a scale to spatial locations remains open. Each object in the image imposes a geography of its own. Multiple objects impose multiple geographies, so the assignment of scale must be based on the geographies. The complication is that of handling points which are influenced by the geography of more than one object. The same complication arises in the analysis of figure–subfigure hierarchies, where the geometry of subfigures is related to the geometry of the figures.

Our multiscale analysis involves real–valued functions of both space and scale, for example, the medialness and boundariness functions. It is my opinion that a better understanding of these functions must be developed by studying the geometry of their graphs. Thus, for example, an investigation is needed to decide on an appropriate metric for the Cartesian product of scale space and the reals. The consequences, such as how to construct ridges on the functions' graphs, must then be studied.