ROSS T. WHITAKER. Geometry-Limited Diffusion (Under the direction of STEPHEN M. PIZER.)

Abstract

This paper addresses the problem of image segmentation and suggests that a number of interesting visual features can be specified as the boundaries of segments. It proposes a segmentation strategy that utilizes the local homogeneity of local image structure. Stable differential measurements are made via a variable-conductance diffusion process, sometimes called anisotropic diffusion. Anisotropic diffusion has been shown to preserve important image structure while reducing unwanted noise.

A multi-scale approach to variable-conductance diffusion is described. This technique combines information at a range of scales and provides a means of obtaining precise boundaries of "large-scale" objects in the presence of noise. These ideas generalize to incorporate multi-valued images and higher-order geometric structure. The result is a framework for constructing image segmentations based on the homogeneity of a set of descriptors. When applied to local image geometry, this method offers a means of breaking images into geometric patches that have interesting visual features as boundaries.

Several examples of this framework are presented. While the diffusion of intensity provides a means of detecting edges, the diffusion of first-order information produces ridges and valleys. A combination of of first- and second-order information provides a means of detecting the "middles" as well as the edges of objects. A proof of the geometric invariance under orthogonal group transformations is given. These same methods generalize to include systems of frequency-sensitive diffusion processes. These systems are useful for characterizing patches based on texture.

An empirical investigation of these nonlinear diffusion processes shows they are stable in the presence of noise and offer distinct advantages over conventional linear smoothing. This analysis derives from a novel method for evaluating low-level computer-vision algorithms. The method uses ensembles of stochastically generated images to compute a set of statistical metrics which quantify the accuracy and reliability of a process. This evaluation offers a means of comparing different algorithms and exploring the space of free parameters associated with a single algorithm.

Acknowledgments

I owe a great debt to Steve Pizer. He has been an excellent teacher, advisor, and friend. He has been generous with his time and enthusiasm, and he has had a large part in making my years at UNC both professionally and personally rewarding. Perhaps the best repayment I can offer Steve is the promise that I will some day try to give my own students those things that he has given me.

Thanks also to the other members of my committee, James Coggins, James Damon, John Halton, and Jonathan Marshall. Each of them has spent a great deal of time reading this document, and each has given me valuable feedback during the various phases of my research.

Several of my peers have been particularly helpful. Thanks to the members of the MIP group for their feedback during group meetings. Thanks to Bryan Morse and Terry Yoo, to whom I have gone for technical advice on numerous occasions, and thanks to Dave Eberly for, among other things, helping me understand tensor notation well enough to complete the proof of geometric invariance.

Thanks to my wife Kerry for her companionship, moral support, and proof reading.

The greatest thanks go to my mother and father for providing me, throughout my life, with the love and guidance that have made me a happy and productive person.

This research has been supported by NIH grant #P01 CA47982 and the UNC Department of Computer Science Alumni Fellowship.

Contents

List of Tables

List of Figures

List of Symbols

∇	"nabla", the gradient operator
δ	"delta", the Laplacian operator
$\emptyset$	"null" symbol, the empty set
$\cap$	"cap", set intersection
$\cup$	"cup", set union
$\mathbb{R}$	the set of real numbers
$\mathbb{I}$	the set of integers
$\mathbb{D}$	the distance transform in the image space
$\mathbb{C}$	the set of complex numbers
$\mapsto$	"maps to"
$\subseteq$	subset
$\subset$	proper subset
$\supseteq$	superset
$\supset$	proper superset
$\otimes$	convolution
∂	partial derivative
$\stackrel{\mathrm{def}}{=}$	defined as
$\vec{}$	used with other symbols indicates a multi-valued or vector-valued quantity
$\mathcal{D}$	"caligraphic D", the dissimilarity measure
$\mathcal{M}$	"caligraphic M", a metric
$\mathcal{F}$	"caligraphic F", the Fourier transform or the set of functions of polynomial invariants
$\mathcal{N}$	"caligraphic N", a neighborhood
$\mathcal{P}$	"caligraphic P", the set of polynomial invariants
$\mathcal{G}$	"caligraphic G", a low-level computer vision algorithm or process
$\sum$	sum

$\int$	integral
$\in$	"element of"
$\notin$	"not an element of"
$\forall$	"for all"
$\exists$	"there exists"
$\Rightarrow$	"implies"
$\Leftrightarrow$	"implies and is implied by"
$\sqrt{}$	square root
$\cdot$	"dot", the dot product of two quantities or a set of unspecified parameters in a function
∞	infinity