

Image Geometry Through Multiscale Statistics

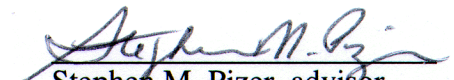
by
Terry Seung-Won Yoo

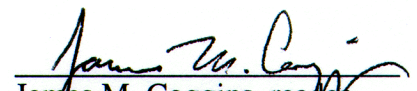
A Dissertation submitted to the faculty of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the
Department of Computer Science.


Chapel Hill

1996

Approved by:


Stephen M. Pizer, advisor


James M. Coggins, reader


J. S. Marron, reader


George D. Stetten, reader

Abstract

This study in the statistics of scale space begins with an analysis of noise propagation of multiscale differential operators for image analysis. It also presents methods for computing multiscale central moments that characterize the probability distribution of local intensities. Directional operators for sampling oriented local central moments are also computed and principal statistical directions extracted, reflecting local image geometry. These multiscale statistical models are generalized for use with multivalued data.

The absolute error in normalized multiscale differential invariants due to spatially uncorrelated noise is shown to vary non-monotonically across order of differentiation. Instead the absolute error decreases between zeroth and first order measurements and increases thereafter with increasing order of differentiation, remaining less than the initial error until the third or fourth order derivatives are taken.

Statistical invariants given by isotropic and directional sampling operators of varying scale are used to generate local central moments of intensity that capture information about the local probability distribution of intensities at a pixel location under an assumption of piecewise ergodicity. Through canonical analysis of a matrix of second moments, directional sampling provides principal statistical directions that reflect local image geometry, and this allows the removal of biases introduced by image structure. Multiscale image statistics can thus be made invariant to spatial rotation and translation as well as linear functions of intensity.

These new methods provide a principled means for processing multivalued images based on normalization by local covariances. They also provide a basis for choosing control parameters in variable conductance diffusion.

Acknowledgements

I am indebted to many people for the successful completion of this document. I am grateful for the generous support of my advisor, Dr. Stephen M. Pizer, who has been with me throughout the years as mentor, colleague, editor, and friend. I also extend special thanks to Dr. James M. Coggins who has been an unending source of advice.

I thank the other members of my committee, Dr. Jonathan A. Marshall, Dr. George D. Stetten, Dr. Benjamin Tsui, and especially Dr. J. S. Marron.

This research has been funded in part through the National Institutes of Health, grant number P01 CA 47982 and the National Science Foundation, NSF ASC-89-20219.

There are many other people without whom I would never have made it to the end of a successful graduate career. Thanks to Murray Anderegg for many cups of coffee and countless sanity-preserving hands of cribbage. Thanks to David Harrison who added breadth to my education in important ways. Thanks to Dr. Richard L. Holloway for the many needed distractions, his sharing of life's little important things, but most of all for his valued friendship. Thanks to my life friends Dr. Don and Claire Stone, without whom I would never have reached the home stretch. Thanks to Matt Fitzgibbon, Dr. Mary McFarlane and Dr. Greg Turk for being there when I've needed them the most.

During my graduate student tenure, I have had many gifted advisors and mentors. For their friendship and guidance I would like to thank Dr. Henry Fuchs, Dr. Frederick Brooks, Kathy Tesh, and Linda Houseman. For their contributions as colleagues, co-authors, conspirators, and confidants I thank David T. Chen, Marc Olano, Steven Aylward, Rob Katz, Dr. Bryan Morse, and Dr. Ross Whitaker.

Finally, I owe my greatest debts to my family. I thank my parents for life and the strength and determination to live it. Special thanks to my son, Ross, who reminds me daily that miracles exist everywhere around us. Most of all, I thank my beloved wife, Penny, who shares my burdens and my joys. This dissertation is equally her achievement. For me, it is she who makes all things possible.

Contents

Abstract.....	i
Acknowledgements	ii
Contents	iv
List of Tables	vii
List of Figures.....	viii
List of Symbols	xi
Chapter 1 Introduction.....	1
1.1. A multiscale approach to computer vision.....	2
1.2. An integrated approach to early vision	3
1.3. Driving issues.....	4
1.4. Thesis	5
1.5. Overview	6
1.6. Contributions	7
Chapter 2 Background	9
2.1. Notation.....	9
2.2. Images	9
2.2.1. Images as a 2D manifold in n-space	10
2.2.2. Digital Images	11
2.3. Invariance.....	12
2.3.1. Gauge Coordinates	13
2.4. Scale Space	14
2.4.1. Differentiation.....	15
2.4.2. The Gaussian as a unique Regular Tempered Distribution.....	16
2.4.3. Zoom Invariance	18
2.4.4. Gaussian Scale Space.....	18
2.5. Image Statistics	19
2.5.1. The Normal Density vs. the Gaussian Filter Kernel	19
2.5.2. Noisy Images.....	20
2.5.3. Statistical Measures as Invariants: Mahalanobis Distances	20
2.5.4. Calculating Central Moments	21
2.5.5. Characteristic Functions.....	23
Univariate Characteristic Functions.....	23
A Simple Univariate Example (a Gaussian Normal Distribution)	23
Bivariate Characteristic Functions	25
2.6. Moment Invariants of Image Functions	26
Chapter 3 Normalized Scale Space Derivatives:	27
3.1. Introduction and Background.....	27
3.1.1. Scale Space Differential Invariants	28
3.1.2. Reconstruction of Sampled Images via the Taylor Expansion	29
3.1.3. Exploring the Properties of Scale-space Derivatives.....	30

3.2. Noise and Scale.....	31
3.3. Variance of Multiscale Derivatives without Normalization	32
3.3.1. Covariances of 1D Multiscale Derivatives	33
3.3.2. Covariances of 2D Multiscale Derivatives	35
3.4. Variance of Normalized Scale Space Derivatives	36
3.5. Analysis of 1D Scale Space Derivatives	37
3.6. Analysis of 2D Scale Space Derivatives	39
3.7. Discussion	42
3.8. Conclusion	43
3.A. Appendix : Covariance of Scale Space Derivatives.....	44
Chapter 4 Multiscale Image Statistics.....	49
4.1. Background and Introduction.....	49
4.2. Images and Stochastic Processes	51
4.2.1. Stochastic Processes.....	51
4.2.2. Images as Samples	52
4.2.3. Ergodicity	53
4.2.4. Ergodicity and Images.....	54
4.3 Multiscale Statistics	55
4.3.1. Multiscale Mean.....	56
4.3.2. Multiscale Variance	57
4.3.3. Multiscale Skewness and Kurtosis.....	59
4.3.4. Invariance with respect to linear functions of intensity	61
4.4. Other Multiscale Central Moments.....	62
4.5. Characteristics of Multiscale Image Statistics	62
4.5.1. Multiscale Statistics vs. Difference of Gaussian Operators	62
4.5.2. Multiscale Moments of Intensity vs. Moment Invariants of Image Functions.....	64
4.6. Measurement Aperture, Object Scale, and Noise	65
4.6.1. Noise Propagation in Multiscale Statistics of an Ergodic Process.....	65
4.6.2. Noise Propagation in Multiscale Statistics of a Piecewise Ergodic Process.....	66
4.7. Multiscale Statistics of 2D Images.....	68
4.7.1. Multiscale 2D Image Mean.....	68
4.7.2. Multiscale 2D Image Variance.....	68
4.7.3. Other Multiscale 2D Image Statistics	69
4.7.4. Some 2D Examples of Multiscale Image Statistics	69
4.8. An Application: Statistical Nonlinear Diffusion.....	70
4.9. Multiscale Statistics of Multivalued Images	72
4.9.1. The Multiscale Multivalued Mean.....	73
4.9.2. Multiscale Multivalued Joint Moments	73
4.9.3. Multiscale Multivalued Variance.....	73
4.9.4. Multiparameter VCD, a foreshadow of future work.....	74
4.10. Summary and Conclusions.....	76
Chapter 5 Directional Multiscale Image Statistics.....	79
5.1. Approaches to Directional Analysis.....	80
5.1.1. Steerable Filters	80

5.1.2. Matrices and Differential Geometry	80
5.1.3. Intensity Invariance vs. Spatially Invariant Directional Analysis	81
5.2. Directional Statistics	82
5.2.1. Multiscale Directional Means	82
5.2.2. Multiscale Directional Covariances	83
5.2.3. The Cauchy-Schwarz Inequality for Multiscale Directional Covariances	85
5.3. Directional Multiscale Statistics of Sample 2D Images.....	87
5.4. The Directional Multiscale Covariance Matrix	89
5.5. SVD Applied to Directional Multiscale Statistics of 2D Images.....	91
5.6. Multiscale Gauge Coordinates of Image Statistics	92
5.7. Invariants of Directional Multiscale Moments	93
5.8. Multiscale Directional Statistics of Multivalued Images.....	95
5.8.1. Canonical Correlation Analysis of Multivalued Directional Statistics.....	96
5.8.2. Understanding Canonical Correlations of Multivalued Directional Statistics	98
5.9. Covariance between Image Intensity and Space	99
5.9.1. Directional Analysis of 2D Scalar Images	99
5.9.2. Canonical Correlation Analysis versus Differential Operators.....	101
5.10. Summary	102
5.A. Appendix: Singular Value Decomposition of a 2x2 Symmetric Matrix.....	103
Chapter 6 Conclusions and Future Directions.....	105
6.1. Contributions and Conclusions.....	106
6.2. Future Directions in Multiscale Statistical Theory	107
6.2.1. Local Differential Geometry and Local Image Statistics	108
6.2.2. Multiscale Distribution Analysis	108
6.2.3. Comparing Two Distributions	109
6.3. Applying Multiscale Image Statistics	119
6.3.1. Statistical Control of Nonlinear Diffusion.....	110
6.3.2. Mixtures in Segmentation.....	111
6.4. Multiscale Image Statistics in Medicine	112
6.5. Summary	113
Bibliography	115

List of Tables

Table 3.1. Variances of unnormalized scale space derivatives (order 0-6) of noisy 1D images (variance of input noise = v_0).....	35
Table 3.2. Variances of unnormalized scale space derivatives of noisy 2D images (variance of input noise = v_0) for partial spatial derivatives to the fourth order (Adapted from Blom 1992).....	35
Table 3.3. Variances of normalized scale space derivatives (order 0-6) of noisy 1D images (variance of input noise = v_0).....	37
Table 3.4. Variances of normalized scale space derivatives of noisy 2D images for partial spatial derivatives to the fourth order (variance of input noise = v_0).....	37
Table 3.5. Variances of both unnormalized and normalized scale space derivatives (order 0-6) of noisy 2D images (variance of input noise = v_0)	41

List of Figures

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.	3
Figure 2.1. Three representations of an image. From left to right: (a) greyscale representation, (b) intensity surface, and (c) isophotes.....	10
Figure 2.2. The image in Figure 2.1 represented as a digital image with a raster resolution of 64×64 pixels.	11
Figure 2.3. 2-D Gaussian derivative filter kernels through the 4th order.	16
Figure 2.4. <i>Top:</i> Characteristic function for a zero mean Gaussian. Maclaurin approximating polynomials (a) $n=2$, (b) $n=8$, (c) $n=10$, and (d) $n=16$	25
Figure 3.1. Propagated error of unnormalized 1D scale space derivatives (order 0-6). Each curve represents the ratio of variance of output to input noise of the linear unnormalized derivative of Gaussian filter vs. scale σ . Plot is on a log-log scale....	38
Figure 3.2 Plot of the propagated error of normalized 1D scale space derivatives (order 0-6). Each curve represents the ratio of variance of output to input noise of the linear normalized derivative of Gaussian filter vs. scale σ . Plot is on a log-log scale.....	38
Figure 3.3. Plot of the propagated error of normalized 1D scale-space derivatives (order 0-6). Curve represents the ratio of variance of output to input noise of the linear unnormalized derivative of Gaussian filter vs. order of differentiation. Plot is on a log scale.	39
Figure 3.4. Propagated error of unnormalized 2D scale space derivatives (order 0-6). Each curve represents the ratio of variance of output to input noise of the linear unnormalized derivative of Gaussian filter vs. scale σ . Plot is on a log-log scale....	42
Figure 3.5 Plot of the propagated error of normalized 2D scale space derivatives (order 0-6). Each curve represents the ratio of variance of output to input noise of the linear normalized derivative of Gaussian filter vs. scale σ	42
Figure 3.6. Plot of the propagated error of normalized 2D scale-space derivatives (order 0-6). Curve represents the ratio of variance of output to input noise of the linear unnormalized derivative of Gaussian filter vs. order of differentiation. Plot is on a log scale.	42
Figure 4.1a. Generic 1D square pulse function $P(d, x)$. Used as the input for generating pulse transfer functions.	56
Figure 4.1b. 1D square pulse functions $P(1, x)$, $P(2, x)$, $P(4, x)$, $P(8, x)$. From left to right: $d = 1$, $d = 2$, $d = 4$, $d = 8$; $\lim_{d \rightarrow 0} P(d, x) = \delta(x)$	57
Figure 4.2. 1D pulse transfer function for the multiscale mean operator $\mu_{P(d, x)}(x \sigma)$. From left to right: $d = 1$, $d = 2$, $d = 4$, $d = 8$. In all images, $\sigma = 1$. The dashed lines represent the input pulse function $P(d, x)$. Note the difference in spatial and intensity ranges in each image.	57
Figure 4.3. 1D pulse transfer function for the multiscale variance operator $\mu_{P(d, x)}^{(2)}(x \sigma)$. From left to right: $d = 1$, $d = 2$, $d = 4$, $d = 8$. In all images, $\sigma = 1$. Note the difference in spatial and intensity ranges in each image.	58

Figure 4.4. Comparison of the 1D pulse transfer function for the multiscale variance operator $\mu_{P(d,x)}^{(2)}(x \sigma)$ to the square multiscale gradient magnitude operator. Top row, $\mu_{P(d,x)}^{(2)}(x \sigma)$. Bottom row: $(\frac{\partial}{\partial x} P(d,x \sigma))^2$. From left to right: $d = 1, d = 2, d = 4, d = 8$. In all images, $\sigma = 1$. Note the difference in spatial and intensity ranges in each image.....	59
Figure 4.5. 1D pulse transfer function of $\mu_{P(d,x)}^{(3)}(x \sigma)$. From left to right: $d = 1, d = 2, d = 4, d = 8$. In all images, $\sigma = 1$. Note the difference in spatial and intensity ranges in each image.....	59
Figure 4.6. Comparison of $\mu_{P(d,x)}^{(3)}(x \sigma)$ to $\frac{\partial}{\partial x} P(d,x \sigma)$. Top row, $\mu_{P(d,x)}^{(3)}(x \sigma)$. Bottom row: $\frac{\partial}{\partial x} P(d,x \sigma)$. From left to right: $d = 1, d = 2, d = 4, d = 8$. In all images, $\sigma = 1$. Note the difference in spatial and intensity ranges in each image.	60
Figure 4.7. 1D pulse transfer function of $\mu_{P(d,x)}^{(4)}(x \sigma)$. From left to right: $d = 1, d = 2, d = 4, d = 8$. In all images, $\sigma = 1$. Note the difference in spatial and intensity ranges in each image.....	61
Figure 4.8. Comparison of $\mu_{P(d,x)}^{(4)}(x \sigma)$ to $(\frac{\partial^2}{\partial x^2} P(d,x \sigma))^2$. Top row, $\mu_{P(d,x)}^{(4)}(x \sigma)$. Bottom row: $(\frac{\partial^2}{\partial x^2} P(d,x \sigma))^2$. From left to right: $d = 1, d = 2, d = 4, d = 8$. In all images, $\sigma = 1$. Note the difference in spatial and intensity ranges in each image. ...	61
Figure 4.9. Comparisons of $\mu_{P(d,x)}^{(2)}(x \sigma)$ with $\text{dog}(P(d,x); \sigma_a, \sigma_b)$. The input function is a pulse $P(d,x)$. In all cases, $d = 1$. From left to right: a. $\mu_{P(d,x)}^{(2)}(x \sigma)$ with $\sigma = 1$, b. $\text{Dog}(P(d,x); \sigma_a, \sigma_b)$ with $\sigma_a = \sigma/\sqrt{2}$, $\sigma_b = 1$, and c. $\text{Dog}(P(d,x); \sigma_a, \sigma_b)$ with $\sigma_a = 0$, $\sigma_b = 1$	64
Figure 4.10. Test function $T(h,x)$	67
Figure 4.11. A 128 x 128 pixel Teardrop with Signal to Noise of 4:1	69
Figure 4.12. Local statistical measure of the teardrop from Fig. 4.11.	69
Figure 4.13. A test object. The figure contains structures at different scale. The raster resolution of the object is 128 x 128 pixels.	72
Figure 4.14. Results from the modified multiscale statistical approach to vcd (left: initial image, right: after 75 iterations of vcd).....	72
Figure 4.15. Early work in statistically driven multivalued vcd. A synthetic multivalued image where the values are subject to significant gaussian white noise and with a strong negative correlation between intensity values. (a) - original two valued input image and its scatterplot histogram. (b) - image after processing with vcd and resulting histogram.....	76
Figure 5.1. - A test image with SNR of 4:1 with a raster resolution of 256×256 pixels. 87	87
Figure 5.2. Directional variances of the objects from Figure 5.1. (From left to right: a: $V_{xx} = \mu_{I,xx}^{(2)}(\mathbf{p} \sigma)$, b: $V_{xy} = \mu_{I,xy}^{(2)}(\mathbf{p} \sigma)$, c: $V_{yy} = \mu_{I,yy}^{(2)}(\mathbf{p} \sigma)$). In all images, a grey value is 0, and $\sigma = 2$ pixels. Bright grey to white indicates positive values, and dark grey to black indicates negative values. Each image uses a left handed coordinate system with the origin in the upper left corner, the x-axis oriented to the right, and the y-axis oriented from the toward the bottom of the page.	88

Figure 5.3. Eigenvalue images of the object from Figure 5.1, computed with a spatial aperture or scale σ of 2 pixels. (From left to right: a: λ_1 , b: λ_2). In both of these images, black is zero and bright indicates positive values.....	91
Figure 5.4. Eigenvector image of the object from Figure 5.1, computed with a spatial aperture or scale σ of 2 pixels. The image reflects only the eigenvector \mathbf{u} in the direction of maximum variance at each pixel; the eigenvector \mathbf{v} in the direction of minimum variance is perpendicular to the vectors shown. The lengths of the vector representations indicate relative magnitude.	92
Figure 5.5. (a) Test figure exhibiting significant directional spatial correlation and (b) the local anisotropy statistic \hat{Q} where $\sigma = 3$. In both images, the raster resolution is 256×256	95
Figure 6.1. A 2D dual-echo MR image of the head with its scatterplot histogram.	112
Figure 6.2. MR image of a shoulder acquired using a surface receiving coil. This may represent the ultimate test for this research.....	114

List of Symbols

Symbols are listed in order of their appearance in the text.

\mathbb{R}	real numbers
\mathbb{R}^n	n-space of real numbers
\rightarrow	maps onto
p	bold designates vector or tensor quantities
$I(\mathbf{p})$	function I of p
\in	is an element of
\mathbb{N}	natural numbers, (i.e., 0, 1, 2, 3, ...)
\mathbb{N}^n	n-space of natural numbers
\subset	is a subset of
<i>L</i>	italics designate set notation
\vec{V}	a vector field V
\bullet	dot or inner product operator
∇I	gradient of I
θ	Greek lower-case theta, used to designate an angular value
\otimes	the convolution operator
$G(\sigma, \mathbf{p})$	Gaussian with spatial scale σ .
\int	integral
$\frac{\partial^n}{\partial x^n}$	n-th partial derivative with respect to x
Σ	Greek upper-case sigma, without subscripts, denotes a covariance matrix
σ	Greek lower-case sigma, used as the scale parameter
e	the transcendental value e, the natural logarithm
π	Greek lower-case pi, the transcendental value, pi
\mathbf{p}^T	transpose of the tensor value p
$\nabla \bullet \mathbf{F}$	“del-dot,” the divergence of F
$I(\mathbf{p} \sigma)$	multiscale measurement: function I(p) at scale σ
\tilde{u}	tilde over a character denotes random variable
μ	Greek lower-case mu, used to designate moments
$\langle \rangle$	the expectation operation
$N_{\mu, \mu^{(2)}}(\tilde{u})$	Standard Normal distribution of \tilde{u} , with mean μ and variance $\mu^{(2)}$.
$\mu^{(n)}$	n-th central moment
$\sum_{j=1}^n f(j)$	summation of $f(0) + f(1) + \dots + f(n)$
i	imaginary value $\sqrt{-1}$
$\overline{M}(u, v)$	the u-v spatial moment of an image function
$L_{x^n y^m}$	multiscale partial derivative $\frac{\partial^n}{\partial x^n} \frac{\partial^m}{\partial y^m} G(\sigma, \mathbf{p}) \otimes I(\mathbf{p})$
$\hat{L}_{x^n y^m}$	normalized multiscale partial derivative $\sigma^{n+m} \frac{\partial^n}{\partial x^n} \frac{\partial^m}{\partial y^m} G(\sigma, \mathbf{p}) \otimes I(\mathbf{p})$
\mathbb{Z}	integers

\forall	for all
$V(\tilde{u})$	variance of \tilde{u}
$M(\tilde{u})$	mean of \tilde{u}
$\text{Cov}(\tilde{u})$	covariance of \tilde{u}
$\prod_{k=1}^n f(k)$	product of $f(0) + f(1) + \dots + f(n)$
\underline{F}	stochastic process F
$f(x) \xrightarrow{a \rightarrow \infty} c$	$f(x)$ approaches c as a goes to infinity
$\lim_{d \rightarrow 0} P(d)$	limit of $P(d)$ as d approaches 0
$\text{erf}(x)$	standard error function: $\text{erf}(x) = \int_{-\infty}^x G(1, \tau) \, d\tau$
$ \mathbf{p} $	norm of \mathbf{p} : $\mathbf{p} \bullet \mathbf{p}$
\perp	perpendicular to
$\max_{-\infty < x < \infty} (S(x))$	global maximum of $S(x)$ over the interval $-\infty < x < \infty$