LBNL-40490
UC-405
Preprint

ERNEST ORLANDO LAWRENCE BERKELEY NATIONAL LABORATORY

Images as Embedded Maps and Minimal Surfaces: Movies, Color, Texture, and Volumetric Medical Images

R. Kimmel, R. Malladi, and N. Sochen

Computing Sciences Directorate

June 1997

Submitted to
IEEE Transactions on Pattern Analysis and Machine Intelligence

REFERENCE COPY
Does Not
Circulate
Bldg. 50 Library - Ref.
Lawrence Berkeley National Laboratory
Copy 1
LBNL-40490

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBNL-40490

IMAGES AS EMBEDDED MAPS AND MINIMAL SURFACES: MOVIES, COLOR, TEXTURE, AND VOLUMETRIC MEDICAL IMAGES*

R. Kimmel, R. Malladi, and N. Sochen
Mathematics Department
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720

June 1997

* This work is supported in part by the Applied Mathematics Subprogram of the Office of Energy Research under DE-AC03 76SFOOO98, ONR grant under NOOO14-96-1-0381, and in part by the National Science Foundation under grant PHY-90-21139.

Images as Embedded Maps and Minimal Surfaces: Movies, Color, Texture, and Volumetric Medical Images *

R. Kimmel,[1] R. Malladi,[1] and N. Sochen[2]

[1] Lawrence Berkeley National Laboratory
and
Department of Mathematics
University of California, Berkeley, CA 94720.

[2]Raymond and Beverly Sackle Faculty of Exact Sciences
Tel-Aviv University, Israel.

Abstract

We present a framework for enhancing images while preserving either the edge or the orientation-dependent texture information present in them. We do this by treating images as manifolds in a feature-space. This geometrical interpretation leads to a natural way for grey level, color, movies, volumetric medical data, and color-texture image enhancement. Following this, we invoke the Polyakov action from high-energy physics, and develop a minimization procedure through a geometric flow. This flow, based on manifold volume minimization yields a natural enhancement procedure. We apply this framework to edge-preserving denoising of grey value and color images, for volumetric medical data, and orientation-preserving flows for grey level and color texture images.

Keywords:*Scale-space, Minimal surfaces, PDE based non-linear image diffusion, Selective smoothing, Color processing, Texture enhancement, Movies and volumetric medical data.*

1 Introduction

In this paper, we present a general framework for processing images of various types like grey scale, color, and those that have orientation-dependent information such as textures. We do this by treating images as embedded maps that flow towards minimal surfaces. In

*This work is supported in part by the Applied Mathematics Subprogram of the Office of Energy Research under DE-AC03-76SFOOO98, ONR grant under NOOO14-96-1-0381, and in part by the National Science Foundation under grant PHY-90-21139.

other words, our view on images is that they are $2D$ or $3D$ manifolds embedded in higher dimensional space; for example a grey-scale image is a surface in (x, y, I) space and a color image is a surface embedded in a $5D$ space, i.e. the (x, y, I^r, I^g, I^b). We then use the Polyakov action, that is a general way of measuring area for a manifold embedded in a given space. The edge-preserving enhancement procedure is a result of minimizing this "action" and is expressed via a geometric flow. Our framework has the following properties: (1) It is the most general way of writing the geometrical scale-space and enhancement algorithms for grey-scale, color, volumetric, time-varying, and texture images, (2) it unifies many existing partial differential equation based schemes for image processing, and (3) the schemes are edge-preserving and hence suitable for segmentation tasks.

The importance of edges in scale space construction is obvious. Our view is consistent with the rest of the vision community in that boundaries between objects should survive as long as possible along the scale space, while homogeneous regions should be simplified and flattened in a more rapid way. On the other hand, we still want to preserve the geometry that results in some interesting non-linear 'scale spaces'. Another important question, for which there are only partial answers, is how to treat multi-valued images. A color image is a good example since one actually talks about 3 images (Red, Green, Blue) that are composed into one. In what follows, we attempt to answer this question.

Texture plays an important role in the understanding process of many images, specially those that involve natural scenes. Therefore, it became an important research subject in the fields of psychophysics and computer vision. The study of texture starts from the pre-image that describes the physics and optics that transforms the $3D$ world into an image, through human perception that starts from the image formation on the retina and tracks its interpretation at the first perception steps in the brain.

Preserving the orientation information while diffusing a given texture image is important in certain cases, say in denoising a fingerprint image. We imagine a procedure that preserves domains of constant/homogeneous texture, enhances the texture in each domain, and thereby enhances the boundaries between neighboring domains with different textures. In this paper, we introduce a geometrical way to improve and enhance texture based images. The geometrical feature enhancement procedure we introduce may serve as a step towards segmentation. Weickert in [35, 36] presents a coherence enhancing flow based on a structure tensor idea. We adopt a similar approach in Section 4 and extend the texture enhancement flows to the multi-channel image case, e.g. color textures.

An alternate way of analyzing textures is to represent a given image as a set of sub-band images using the $2D$ Gabor/Morlet-wavelet transform. Some nice mathematical properties and the relation of this transform to the physiological behavior were studied in [17, 27]. This model was later used for the segmentation, interpretation and analysis of texture [4, 18], and for texture-based browsing [21]. In Section 5, we use the Gabor/Morlet-wavelet transform to split a given image into a set of sub-band images. We then show that an enhancement procedure can be constructed based on a flow in the transformed space, i.e. the transform coefficients are treated as higher dimensional manifolds.

The remainder of this paper is organized as follows: Section 2 reviews the definition of arclength, the consideration of images as surfaces, and the minimization of Polyakov action, that leads to a geometric flow that we named *Beltrami flow*. Next, in Section 3 we present

the metric and the resulting edge-enhancing flow for color images. In Section 4, together with some texture enhancement results, we describe an orientation diffusion procedure in color that is deduced from the Beltrami flow. In Section 5 we present an alternate way of texture analysis using $2D$ Gabor/Morlet-wavelet transform. Finally, the analysis of movies and volumetric medical images is presented in Section 6.

2 Images as Embedded Maps that flow toward Harmonic Maps

We consider images as surfaces in higher dimensional spaces [34], and construct enhancement procedures for color images as 2D surfaces in 5D (x, y, I^r, I^g, I^b) space. The idea of images as curved spaces is not limited to 2D surfaces; movies and volumetric images can be considered as 3D hypersurfaces in 4D, e.g. (x, y, z, I) space [15].

Our geometric framework finds a seamless link between the L_1 (Rudin Osher Fatemi [30] TV and its variants) and the L_2 norms (used in Mumford-Shah [23] and its variants) based on the geometry of the image and its interpretation as a surface[1]. The aspect ratio between the gray level and the xy plane, used as a parameter, enables us to switch between the two commonly used norms. This observation made it possible to show that our multi-channel (color) enhancement procedure may be considered as a generalization of the powerful TV scheme that is now commonly used in the high tech. image processing industry. This procedure yields promising results for color image enhancement [34].

In this work, we also propose a flow in a rich feature space which is different from the image space. Other flows in similar feature spaces were recently proposed in [31, 29, 7, 33, 37]; see also [35, 36] for orientation preserving flows. All these approaches begin with a flat metric [10] that does not yield a meaningful minimization process when going to more than one channel[2]. The main difference between these schemes and the one we propose is the geometric interpretation of the information as a manifold flowing so as to minimize its volume. Our geometric perspective of a color image as a surface embedded in a higher dimensional space enabled us to define a simple and natural coupling in the multi-channel color space. Other schemes have also considered image as a surface [2, 12, 38, 20], some even used the image information to build a Riemannian metric for segmentation [5]. However, these methods were not generalized to feature space or any co-dimension higher than one. We now describe the details of our framework.

2.1 The Metric

The basic concept of Riemannian differential geometry is distance. Let us start with the map $\mathbf{X} : \Sigma \to \mathbb{R}^3$, where Σ is a $2D$ manifold. We denote the local coordinates on the two dimensional manifold Σ by (σ^1, σ^2). The map X is explicitly given by

$$(X^1(\sigma^1, \sigma^2), X^2(\sigma^1, \sigma^2), X^3(\sigma^1, \sigma^2)).$$

[1]TV (Total Variation) schemes are based on minimizing the L_1 norm, namely $\int |\nabla I|$, the L_2 norm minimizes $\int |\nabla I|^2$, while the area of the gray level image surface is given by $\int \sqrt{1 + |\nabla I|^2}$.

[2]This flat metric is called 'structure tensor' in [35, 36].

Since the local coordinates σ^i are curvilinear, and not orthogonal in general, the distance square between two close points on Σ, $p = (\sigma^1, \sigma^2)$ and $p + (d\sigma^1, d\sigma^2)$ is not $ds^2 = d\sigma_1^2 + d\sigma_2^2$. In fact, the squared distance is given by a positive definite symmetric bilinear form called the metric, whose components we denote by $g_{\mu\nu}(\sigma^1, \sigma^2)$

$$ds^2 = g_{\mu\nu} d\sigma^\mu d\sigma^\nu = g_{11}(d\sigma^1)^2 + 2g_{12} d\sigma^1 d\sigma^2 + g_{22}(d\sigma^2)^2, \tag{1}$$

where we used Einstein summation convention in the second equality; identical indices that appear one up and one down are summed over. We will denote the inverse of the metric by $g^{\mu\nu}$, so that $g^{\mu\nu} g_{\nu\gamma} = \delta_{\mu\gamma}$, where $\delta_{\mu\gamma}$ is the Kronecker delta.

2.2 Induced Metric

Let $\mathbf{X} : \Sigma \to M$ be an embedding of (Σ, g) in (M, h), where Σ and M are Riemannian manifolds and g and h are their metrics respectively. We can use the knowledge of the metric on M and the map $\mathbf{X}$ to construct the metric on Σ. This procedure, which is denoted formally as $(g_{\mu\nu})_\Sigma = \mathbf{X}^*(h_{ij})_M$, is called the *pullback* for obvious reasons and is given explicitly as follow:

$$g_{\mu\nu}(\sigma^1, \sigma^2) = h_{ij}(\mathbf{X}) \partial_\mu X^i \partial_\nu X^j, \tag{2}$$

where $i, j = 1, ..., \dim M$ are being summed over, and in short we have used $\partial_\mu X^i \equiv \partial X^i(\sigma^1, \sigma^2)/\partial\sigma^\mu$.

We will use the following simple and useful example that is often used in computer vision: Consider embedding of a surface described as a graph in $\mathbb{R}^3$,

$$\mathbf{X} : (\sigma^1, \sigma^2) \to (\sigma^1, \sigma^2, I(\sigma^1, \sigma^2)). \tag{3}$$

Using Eq. (2) we get

$$(g_{\mu\nu}) = \begin{pmatrix} 1 + I_x^2 & I_x I_y \\ I_x I_y & 1 + I_y^2 \end{pmatrix}, \tag{4}$$

where we used the identification $X^1 \equiv \sigma^1$ and $X^2 \equiv \sigma^2$ in the map $\mathbf{X}$.

Actually we can understand this result in an intuitive way: Eq. (2) means that the distance measured on the surface by the local coordinates is equal to the distance measured in the embedding coordinates. Under the above identification, we can write

$$\begin{aligned} ds^2 &= dx^2 + dy^2 + dI^2 \\ &= dx^2 + dy^2 + (I_x dx + I_y dy)^2 \\ &= (1 + I_x^2) dx^2 + 2 I_x I_y dx dy + (1 + I_y^2) dy^2. \end{aligned}$$

2.3 Polyakov Action

Let us briefly review our framework for non-linear diffusion in computer vision. The equations are derived by a minimization problem from an action functional. The functional in question depends on *both* the image manifold and the embedding space. Denote by (Σ, g) the image

manifold and its metric and by (M, h) the space-feature manifold and its metric, then the map $\mathbf{X} : \Sigma \to M$ has the following weight

$$S[X^i, g_{\mu\nu}, h_{ij}] = \int d^m\sigma\sqrt{g}g^{\mu\nu}\partial_\mu X^i \partial_\nu X^j h_{ij}(\mathbf{X}), \tag{5}$$

where m is the dimension of Σ, g is the determinant of the image metric, $(g^{\mu\nu})$ is the inverse of the image metric, the range of indices is $\mu, \nu = 1, \ldots, \dim\Sigma$, and $i, j = 1, \ldots, \dim M$, and (h_{ij}) is the metric of the embedding space[3]. This functional, for $m = 2$, was first proposed by Polyakov [26] in the context of high energy physics.

Given the above functional, we have to choose the minimization. We may choose for example to minimize with respect to the embedding alone. In this case the metric $(g_{\mu\nu})$ is treated as a parameter and may be fixed by hand. Another choice is to vary only with respect to the feature coordinates of the embedding space, or we may choose to vary the image metric as well. In [34] we show how different choices yield different flows. Some flows are recognized as existing methods, other choices are new and will be described below.

Using standard methods in variational calculus (see [34]), the Euler-Lagrange equations with respect to the embedding are:

$$-\frac{1}{2\sqrt{g}}h^{il}\frac{\delta S}{\delta X^l} = \frac{1}{\sqrt{g}}\partial_\mu(\sqrt{g}g^{\mu\nu}\partial_\nu X^i). \tag{6}$$

Our proposal is to view scale-space as the gradient descent:

$$X_t^i \equiv \frac{\partial X^i}{\partial t} = -\frac{1}{2\sqrt{g}}h^{il}\frac{\delta S}{\delta X^l}. \tag{7}$$

Notice that we used our freedom to multiply the Euler-Lagrange equations by a strictly positive function; since $(g_{\mu\nu})$ is positive definite, $g \equiv \det(g_{\mu\nu}) > 0$ for all σ^μ. This factor is the simplest one that does not change the minimization solution while giving a reparametrization invariant expression. The operator that is acting on X^i is the natural generalization of the Laplacian from flat spaces to manifolds and is called *the second order differential parameter of Beltrami* [16], or for short *Beltrami operator*, and is denoted by Δ_g.

For a surface Σ, embedded in 3 dimensional Euclidean space, we get a minimal surface as the solution to the minimization problem. In order to see that and to connect to the usual representation of the minimal surface equation, we notice that the solution of the minimization problem with respect to the metric is

$$g_{\mu\nu} = \partial_\mu X^i \partial_\nu X_i. \tag{8}$$

On inspection, this equation is simply the induced metric on Σ. For the case of a surface embedded in $\mathbb{R}^3$ we calculated it explicitly in (see Eq. (4)). Plugging this induced metric in the first Euler-Lagrange, Eq. (6), we get the steepest decent flow

$$\mathbf{X}_t = H\vec{\mathcal{N}}, \tag{9}$$

[3]E.g. $g^{\mu\nu}$ are the elements of the matrix $(g^{\mu\nu}) = \begin{pmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{pmatrix} = g^{-1}\begin{pmatrix} g_{22} & -g_{12} \\ -g_{12} & g_{11} \end{pmatrix}$.

where H is the mean curvature, $\vec{\mathcal{N}}$ is the normal to the surface[4]:

$$\begin{aligned} H &= \frac{(1+I_x^2)I_{yy} - 2I_xI_yI_{xy} + (1+I_y^2)I_{xx}}{g^{\frac{3}{2}}}, \\ \vec{\mathcal{N}} &= \frac{1}{\sqrt{g}}(-I_x, -I_y, 1)^T, \end{aligned} \tag{10}$$

and $g = 1 + I_x^2 + I_y^2$. Note that this choice gives us the mean curvature flow which means that every point of the surface moves in the normal direction at a speed that is equal to the mean curvature. This should not be a surprise, since the action functional for the above choice of metric $g_{\mu\nu}$ is

$$S = \int d^2\sigma\sqrt{g} = \int d^2\sigma\sqrt{\det(\partial_\mu X^i \partial_\nu X_i)},$$

which is the Euler functional that describes the area of the surface (also known in high energy physics as the Nambu action).

In general, for any manifold Σ and M, the map $\mathbf{X} : \Sigma \to M$ that minimizes the action S with respect to the embedding is called a *harmonic map.* The harmonic map is the natural generalization of the geodesic curve and the minimal surface to higher dimensional manifolds and for different embedding spaces.

The generalization to any manifold embedded with arbitrary co-dimension is given by using Eq. 6 for all the embedding coordinates and the induced metric Eq. 8; see [34] for more details.

2.4 The Beltrami Flow

We now present a new and natural flow for images as surfaces. First let us consider the case in which the gray level image is regarded as an embedding map $\mathbf{X} : \Sigma \to \mathbb{R}^3$, where Σ is a two dimensional manifold, and the flow that we derive is natural in the sense that it minimizes the action functional with respect to I and $(g_{\mu\nu})$, while being reparametrization invariant. The coordinates X^1 and X^2 are parameters from this view point and are identified as above with σ^1 and σ^2 respectively. The result of the minimization is the Beltrami operator acting on I:

$$I_t = \Delta_g I \equiv \frac{1}{\sqrt{g}}\partial_\mu(\sqrt{g}g^{\mu\nu}\partial_\nu I) = \langle \vec{\mathcal{N}}, \hat{I}\rangle H, \tag{11}$$

where the metric is the induced one given in Eq. 4, $\hat{I}$ is the unit vector in the I direction, and $\langle \vec{\mathcal{N}}, \hat{I}\rangle$ is the $\hat{I}$ component of the image surface normal $\vec{\mathcal{N}}$ (i.e. projection of $\vec{\mathcal{N}}$ onto $\hat{I}$). See [39], for a recent related effort.

The geometrical meaning of this flow is the following: Each point on the image surface moves with a velocity that depends on the mean curvature vector $H\vec{\mathcal{N}}$ projected to the $\hat{I}$ direction at that point. Since along the edges the normal to the surface lie almost entirely in the x-y plane, I hardly changes along the edges, while the flow drives other regions of the

[4]Some definitions of the mean curvature include a factor of 2 that we omit here.

image towards a minimal surface at a more rapid rate. In other words, this flow preserves edges. As an example, Fig. 1 shows the result of the Beltrami flow applied to a computed tomography (CT) image of the abdominal section. It demonstrates the edge preserving property of the Beltrami flow.

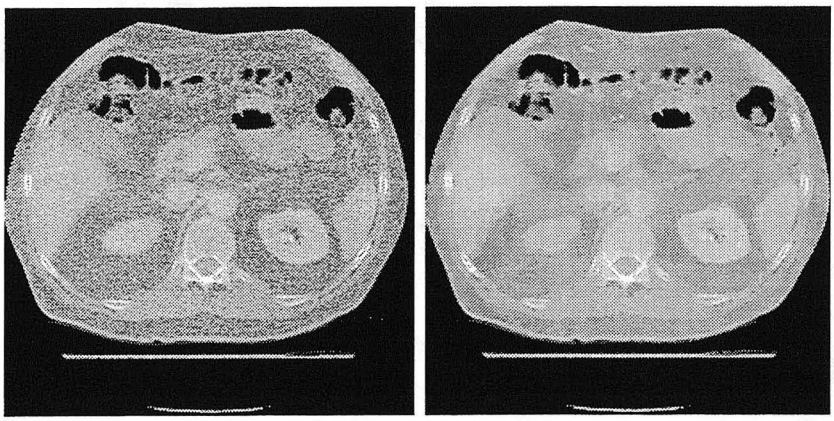

Figure 1: Left: Original medical image. Right: Result of the Beltrami flow.

In what follows we apply this operator to construct an orientation-preserving flow on texture images. But first let us look at the color image case more closely.

3 Color

We generalize the Beltrami flow to the 5 dimensional space-feature needed in color images. The embedding space-feature space is taken to be Euclidean with Cartesian coordinate system. The image, thus, is the map $f : \Sigma \to \mathbb{R}^5$ where Σ is a two dimensional manifold. Explicitly the map is

$$f = \left(X^1(\sigma^1,\sigma^2), X^2(\sigma^1,\sigma^2), I^r(\sigma^1,\sigma^2), I^g(\sigma^1,\sigma^2), I^b(\sigma^1,\sigma^2)\right).$$

Note that there are obvious better selections to color space definition rather than the RGB flat space.

We minimize our action (5) with respect to the metric and with respect to (I^r, I^g, I^b). For convenience we denote below (r, g, b) by $(1, 2, 3)$, or in general notation i. Minimizing the metric gives, as usual, the induced metric which is given in this case as follows:

$$\begin{aligned} g_{11} &= 1 + (I^1_x)^2 + (I^2_x)^2 + (I^3_x)^2, \\ g_{12} &= I^1_x I^1_y + I^2_x I^2_y + I^3_x I^3_y, \\ g_{22} &= 1 + (I^1_y)^2 + (I^2_y)^2 + (I^3_y)^2, \\ g &= \det(g_{\mu\nu}) = g_{11}g_{22} - g_{12}^2. \end{aligned} \tag{12}$$

Note that this metric differs from the Di Zenzo matrix [10] that was proposed for the multichannel case (which is not a metric) by the addition of identity matrix (adding 1 to g_{11} and

g_{22}). The source of the difference is the map used to describe the image. Di Zenzo used $\mathbf{X} : \mathbb{R}^2 \to \mathbb{R}^3$ while we use $\mathbf{X} : \Sigma \to \mathbb{R}^5$.

The action functional under this choice of the metric is the Euler functional $S = \int d^2\sigma\sqrt{g}$. It is simply the area of the image surface. Minimization with respect to I^i gives the Beltrami flow

$$I_t^i = \frac{1}{\sqrt{g}}\partial_\mu(\sqrt{g}g^{\mu\nu}\partial_\nu I^i), \tag{13}$$

which is a flow towards a minimal surface that preserves edges.

For simple implementation of the Beltrami flow, we first compute the matrices: I_x^i, I_y^i, and 6 other matrices, namely

$$\begin{aligned} p^i &= g_{22}I_x^i - g_{12}I_y^i, \\ q^i &= -g_{12}I_x^i + g_{11}I_y^i. \end{aligned} \tag{14}$$

Then the evolution is given by

$$I_t^i = \frac{1}{g}\left(p_x^i + q_y^i\right) - \frac{1}{2g^2}\left(g_x p^i + g_y q^i\right), \tag{15}$$

where $g_x = \partial_x g$ ($g_y = \partial_y g$).

A detailed justification for the color norm is given in [14]. Intuitively, for most color image formation models, the gradient directions of the different channels should align. The Beltrami flow un-twists the undesired torsion between the channels while smoothing them and preserving the edges. It is thus considered as a proper candidate for color processing.

3.1 Color Enhancement Results

We now present some results of denoising color images using our model. Spatial derivatives are approximated using central differences and an explicit Euler step is employed to reach the solution. We represent the image in the RGB space; however, other representations and different numerical schemes (as in [6]) are possible.

The results are presented in Fig. 2 as follows: The first shows denoising of a color image corrupted with Gaussian noise; left image is the noisy one and the reconstruction result by applying Beltrami flow is shown on the right. Iteration has been manually stopped to produce the result. Constraints similar to [30, 3] can be added; see [34] for details. The left image on the second row shows a noisy woman's profile. No artificial noise has been added in this case. The enhanced image is shown on the right.

Finally, the third row of the figure presents the result of applying the Beltrami flow to reconstruct a color image with noise artifacts introduced by JPEG lossy compression algorithm. Again, the left image depicts the corrupted one and the right image is the reconstructed one using Beltrami flow.

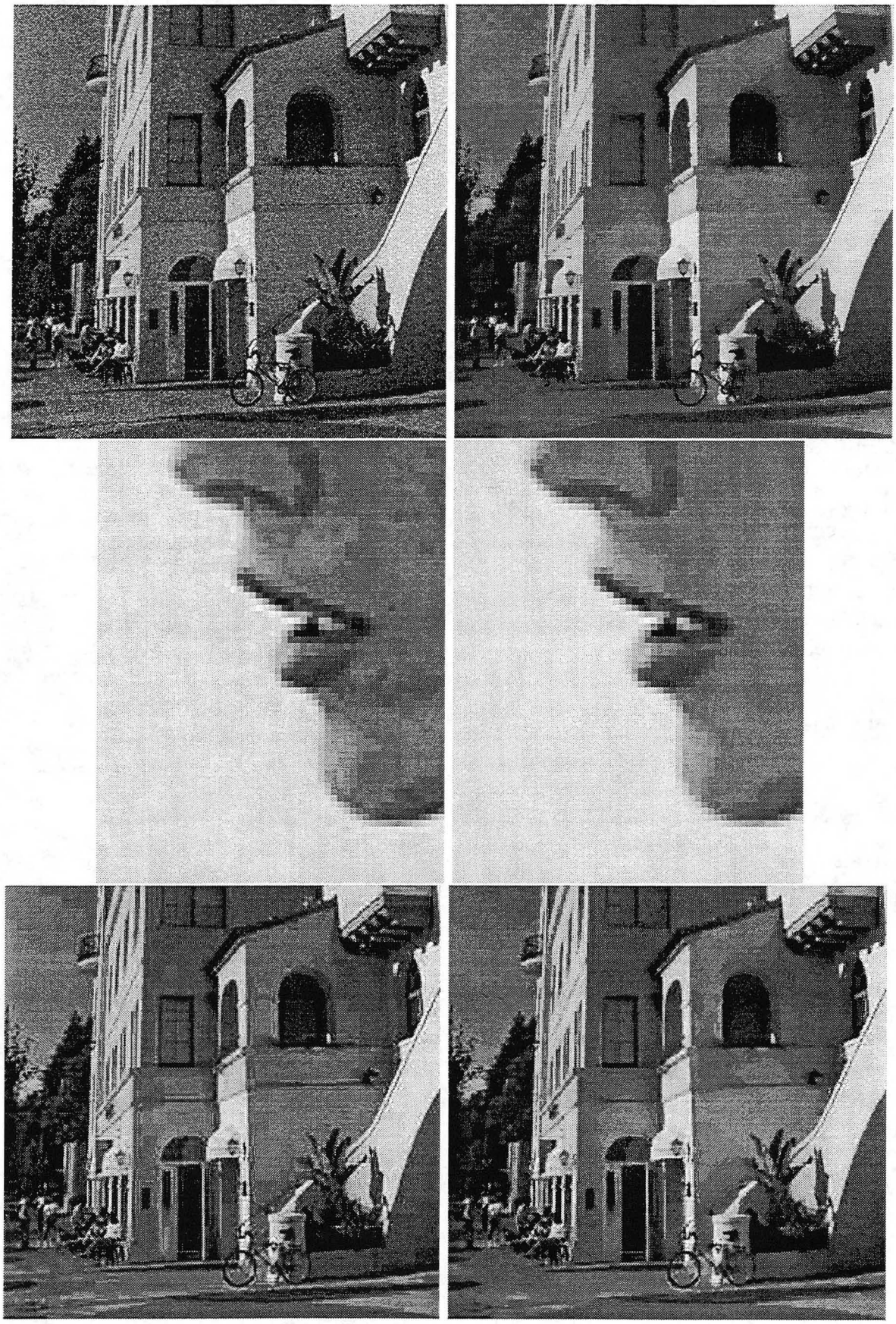

Figure 2: Color results.

4 The Metric as a Structure Tensor

In [13, 19], Gabor considered an image enhancement procedure based on a single small time step along a directional flow. It is based on the anisotropic flow via the inverse second directional derivative in the 'edge' direction (∇I direction) and the geometric heat equation (second derivative in the direction parallel to the edge). The same idea of steering the diffusion direction motivated many recent works[5]. Cottet and Germain [8] used a smoothed version of the image to direct the diffusion, while Weickert [35, 36] smoothed also the structure tensor $\nabla I \nabla I^T$ and then manipulated its eigenvalues to steer the smoothing direction. Eliminating one eigenvalue from a structure tensor proposed in [10] (without smoothing its coefficients) was extended to color space in [33, 32], in which the tensors are not necessarily positive definite. See also [7], where the diffusion is in the direction perpendicular to the maximal gradient of the three channels (this direction is different than that of [33]).

Motivated by all of these results we now present a multi channel extension to Weickert [35] gray level anisotropic orientation diffusion. We show that the diffusion directions can be deduced from the smoothed metric coefficients $g_{\mu\nu}$ and may thus be included within the Beltrami framework under the right choice of directional diffusion coefficients. Based on this observation we now extend the scheme to color and texture.

The induced metric $(g_{\mu\nu})$ is a symmetric uniformly positive definite matrix that captures the geometry of the image surface. Let λ_1 and λ_2 be the largest and the smallest eigenvalues of $(g_{\mu\nu})$, respectively. Since $(g_{\mu\nu})$ is a symmetric positive matrix its corresponding eigenvectors u_1 and u_2 can be chosen orthonormal. Let $U \equiv (u_1|u_2)$, and $\Lambda \equiv \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, then we readily have the equality

$$(g_{\mu\nu}) = U\Lambda U^T. \tag{16}$$

Note also that

$$(g^{\mu\nu}) \equiv (g_{\mu\nu})^{-1} = U\Lambda^{-1}U^T = U \begin{pmatrix} 1/\lambda_1 & 0 \\ 0 & 1/\lambda_2 \end{pmatrix} U^T, \tag{17}$$

and that

$$g \equiv \det(g_{\mu\nu}) = \lambda_1\lambda_2. \tag{18}$$

We will use the image metric in its natural geometric interpretation, i.e. as a structure tensor. The coherence enhancement Beltrami flow $\mathbf{I}_t = \Delta_{\hat{g}}\mathbf{I}$ for color-texture images is then given as follows:

1. Compute the metric coefficients $g_{\mu\nu}$. For the N channel case (for color $N = 3$) we have (see Eq. (12))

$$g_{\mu\nu} = \delta_{\mu\nu} + \sum_{k=1}^{N} I_\mu^k I_\nu^k. \tag{19}$$

[5]This definition of anisotropic flow differs from the Perona-Malik [25] framework, that is locally isotropic. See [28] for many interesting extensions and applications of the locally isotropic flow.

2. Diffuse the $g_{\mu\nu}$ coefficients by convolving with a Gaussian of variance ρ, thereby

$$\tilde{g}_{\mu\nu} = G_\rho * g_{\mu\nu}. \tag{20}$$

For $2D$ images $G_\rho = e^{-(x^2+y^2)/\rho^2}$.

3. Change the eigenvalues, λ_1, λ_2, $\lambda_1 > \lambda_2$, of $(\tilde{g}_{\mu\nu})$ so that $\lambda_1 = \alpha^{-1}$ and $\lambda_2 = \alpha$, for some given positive scalar $\alpha \ll 1$. This yields a new metric $\hat{g}_{\mu\nu}$ that is given by:

$$(\hat{g}_{\mu\nu}) = \tilde{U} \begin{pmatrix} \alpha^{-1} & 0 \\ 0 & \alpha \end{pmatrix} \tilde{U}^T = \tilde{U}\Lambda_\alpha \tilde{U}^T. \tag{21}$$

4. Evolve the k-th channel via Beltrami flow, that by the selection $\hat{g} \equiv \det(\hat{g}_{\mu\nu}) = \lambda_1\lambda_2 = \alpha^{-1}\alpha = 1$ now reads

$$\begin{aligned} I_t^k &= \Delta_{\hat{g}} I^k \equiv \frac{1}{\sqrt{\hat{g}}}\partial_\mu \sqrt{\hat{g}}\hat{g}^{\mu\nu}\partial_\nu I^k = \partial_\mu \hat{g}^{\mu\nu}\partial_\nu I^k \\ &= \operatorname{div}\left(\tilde{U}\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix}\tilde{U}^T \nabla I^k\right) = \operatorname{div}\left(\tilde{U}\Lambda_\alpha \tilde{U}^T \nabla I^k\right). \end{aligned} \tag{22}$$

Note again that for gray level images the above flow is similar to Weickert coherence-enhancing anisotropic diffusion [35, 36] with the important property of a uniformly positive definite diffusion tensor. For gray level images, $(g_{\mu\nu}) = \mathcal{I} + \nabla I \nabla I^T$, where $\mathcal{I}$ is the identity matrix. In this case all that is done is the identity added to Weickert's structure tensor $\nabla I \nabla I^T$. This addition does not change the eigenvectors and thus the proposed flow is equivalent to Weickert scheme. A minor difference is the fact that we propose different manipulation of the eigenvalues. Next we present results of this flow as a filter for orientation enhancing in color images.

4.1 Color Orientation-Enhancing Results

For completeness of the exposition we first repeat the gray level case as in [35, 36] and present an example of a fingerprint enhancement in gray level in Fig. 3.

Next, Fig. 4 presents three cases of color and texture. In the first example we enhance the orientation texture while smoothing the image and reducing the noise. In the next two examples we again adopt Weickert [35] gray level examples and apply the flow, now for more iterations, to two color paintings by van Gogh.

4.2 Beyond a Metric: Reaction in the Edge Direction

Let us take one step further, and exit our 'metric' framework by defining $(g_{\mu\nu})$ to be a non-singular symmetric matrix with one positive and one negative eigenvalues. That is, instead of a small diffusion we introduce a controlled reaction in the edge direction. Here we extend Gabor's idea [13, 19] of inverting the diffusion along the gradient direction.

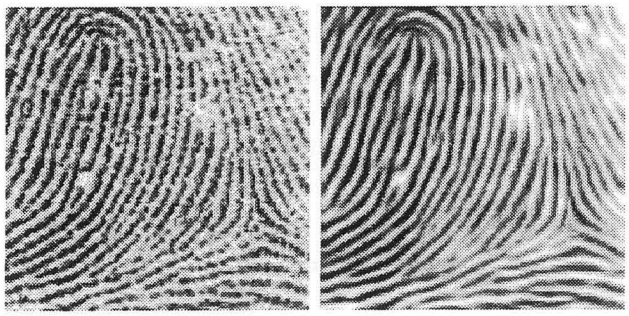

Figure 3: Left: Original fingerprint image 128×128. Right: Result of Beltrami flow with smoothed metric ($\rho = 6$) and steered eigenvalues ($\alpha = 10^{-5}$), after 4 numerical iterations.

Inverting the heat equation is an inherently unstable process. However, if we keep smoothing the metric coefficients, and apply the evolution in the perpendicular direction we get a coherence-enhancing flow with sharper edges that is stable for a short duration of time.

The idea is simply to change the sign of one of the modified eigenvalues in the algorithm described in the previous subsection. We change steps 3 and 4 of the previous scheme that now reads:

1. Compute the metric coefficients $g_{\mu\nu} = \delta_{\mu\nu} + \sum_{k=1}^{N} I_\mu^k I_\nu^k$.

2. Diffuse the $g_{\mu\nu}$ coefficients by convolving with a Gaussian of variance ρ.

3. Change the eigenvalues of $(\tilde{g}_{\mu\nu})$ such that the largest eigenvalue λ_1 is now $\lambda_1 = -\alpha^{-1}$ and $\lambda_2 = \alpha$, for some given positive scalar $\alpha < 1$. This yields a new <u>matrix</u> $\hat{g}_{\mu\nu}$ that is given by:

$$(\hat{g}_{\mu\nu}) = \tilde{U} \begin{pmatrix} -\alpha^{-1} & 0 \\ 0 & \alpha \end{pmatrix} \tilde{U}^T = \tilde{U} \Lambda_\alpha \tilde{U}^T. \tag{23}$$

4. Evolve the k-th channel via the flow, that by the selection $|\hat{g}| \equiv |\det(\hat{g}_{\mu\nu})| = |\lambda_1 \lambda_2| = |-\alpha^{-1}\alpha| = 1$, reads

$$\begin{aligned} I_t^k = \frac{1}{\sqrt{|\hat{g}|}} \partial_\mu \sqrt{|\hat{g}|} \hat{g}^{\mu\nu} \partial_\nu I^k &= \partial_\mu \hat{g}^{\mu\nu} \partial_\nu I^k \\ &= \operatorname{div}\left(\tilde{U} \begin{pmatrix} -\alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} \tilde{U}^T \nabla I^k \right). \end{aligned} \tag{24}$$

For the gray level case with $\rho = 0$ it simplifies to highly unstable inverse heat equation. However, as ρ increases the smoothing along the edges becomes fundamental and the scheme is similar in its spirit to that of [13]. Gabor's [13] comment on the reaction operation in the gradient direction is that *'It is very similar to the operation which the human eye carries*

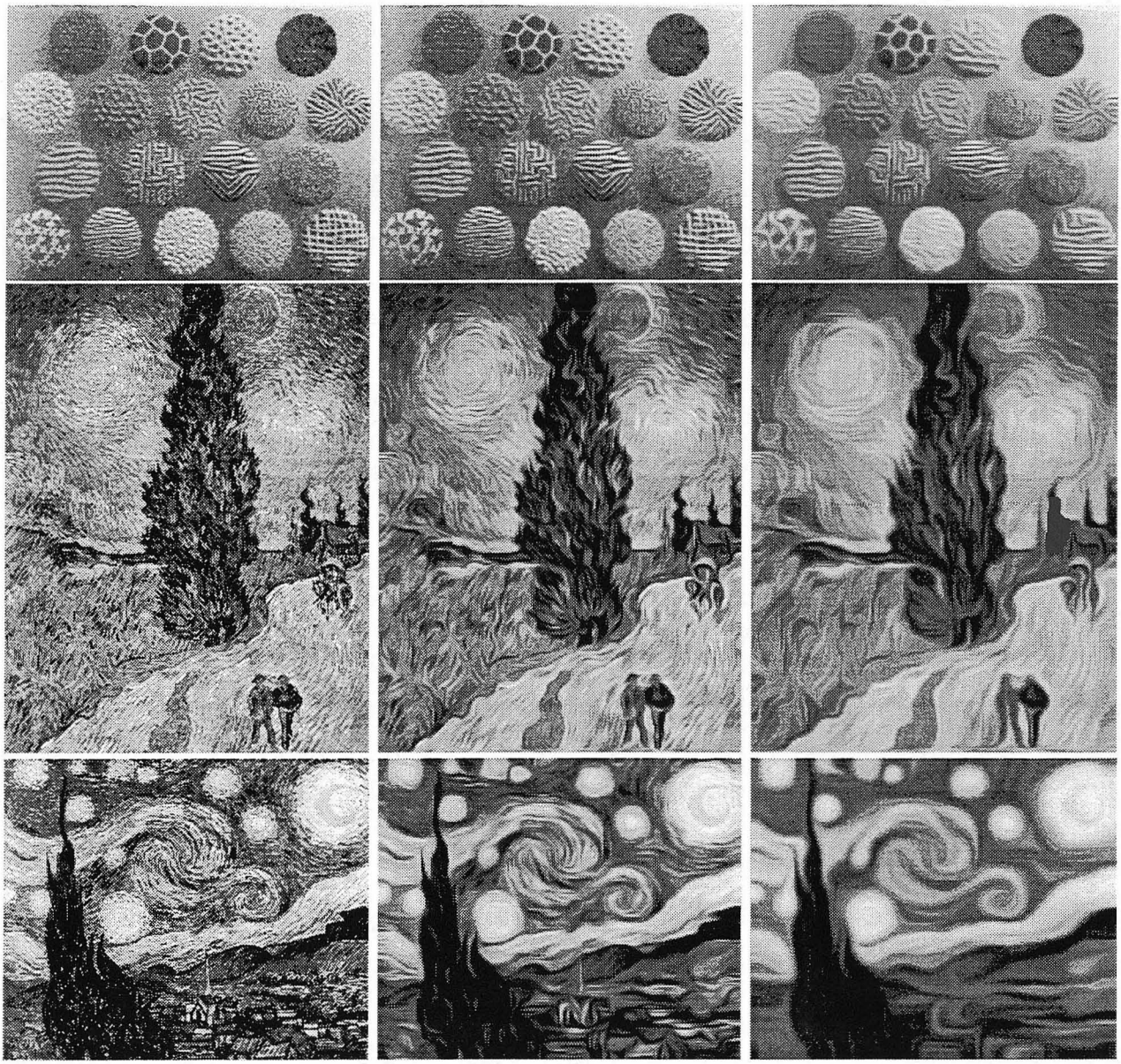

Figure 4: Color and texture: Beltrami flow with smoothed metric and steered eigenvalues ($\alpha = 10^{-5}$); Top Row: Original 'Shells' image 242×184 (left), and the result of the flow ($\rho = 2$) for 2 (middle) and 16 (right) numerical iterations. Middle Row: Original image of van Gogh "Lane under Cypresses below the Starry Sky" 392×512 (left), and the result of the flow ($\rho = 4$) for 16 (middle) and 128 (right) iterations. Bottom Row: Original image of van Gogh "Starry Night" 290×241 (left), and the result of the flow ($\rho = 4$) for 16 (middle) and 256 (right) iterations.

out automatically, and it is not surprising that even the first steps in imitating the human eye by mechanical means lead to rather complicated operations'. It is important to note that the idea of stabilizing the inverse heat equation is extensively used in image processing. Exploring this area is beyond the scope of this paper. However, we like to refer the reader to the 'shock filters' introduced by Osher and Rudin in [24] for gray level images, and the extension of Alvarez and Mazora [1] who apply a reaction in the gradient direction combined with a directional smoothing in the orthogonal direction for gray level images.

Let us apply the above reaction-diffusion algorithm to color images with $\rho = 2$. Figure 5 presents two examples of the flow with a reaction in the edge direction. For comparison to the previous orientation smoothing algorithm, the third row presents two steps along the color orientation diffusion flow.

5 2D Gabor/Morlet-Wavelets as a Space for Texture Images

At the risk of introducing additional computational burden, in this section, we introduce an alternate way of dealing with texture images. In [17] Lee argues that the 2D Gabor/Morlet wavelet transform with specific coefficients is an appropriate mathematical description for images. He motivated his model by recent neurophysiological evidence based on experiments on the visual cortex of mammalian brains. These experiments indicate that a good model for the filter response of simple cells are self-similar 2D Gabor/Morlet wavelets. We refer the interested reader to [22] for implementation considerations, and to the rich literature on wavelet theory, e.g. [9]. Here, we will comment on the basic concepts that are relevant to our discussion.

Following Lee [17], we briefly describe the 2D Gabor/Morlet wavelets that model the simple cells while satisfying Daubechies' wavelet theory [9]. The 2D wavelet transform on an image $I(x, y)$, is defined as

$$(T^{wav}I)(x_0, y_0, \theta, a) = \|a\|^{-1} \int\int dxdy I(x,y)\psi_\theta\left(\frac{x-x_0}{a}, \frac{y-y_0}{a}\right), \tag{25}$$

where a is a dilation parameter, x_0 and y_0 are the spatial translations, and θ is the wavelet orientation parameter.

$$\psi(x, y, x_0, y_0, \theta, a) = \|a\|^{-1}\psi_\theta(\frac{x-x_0}{a}, \frac{y-y_0}{a}), \tag{26}$$

is the 2D elementary wavelet function rotated by θ. Based on neurophysiological experiments, a specific Gabor elementary function is used as the *mother wavelet* to generate the 2D Gabor/Morlet wavelet family by convolving the image with

$$\psi(x, y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{8}(4x^2+y^2)}(e^{ikx} - e^{-\frac{k^2}{2}}), \tag{27}$$

and $\psi_\theta(x, y) = \psi(\tilde{x}, \tilde{y})$ is defined by rotation of (x, y) via

$$\begin{cases} \tilde{x} &= x\cos\theta + y\sin\theta \\ \tilde{y} &= -x\sin\theta + y\cos\theta. \end{cases} \tag{28}$$

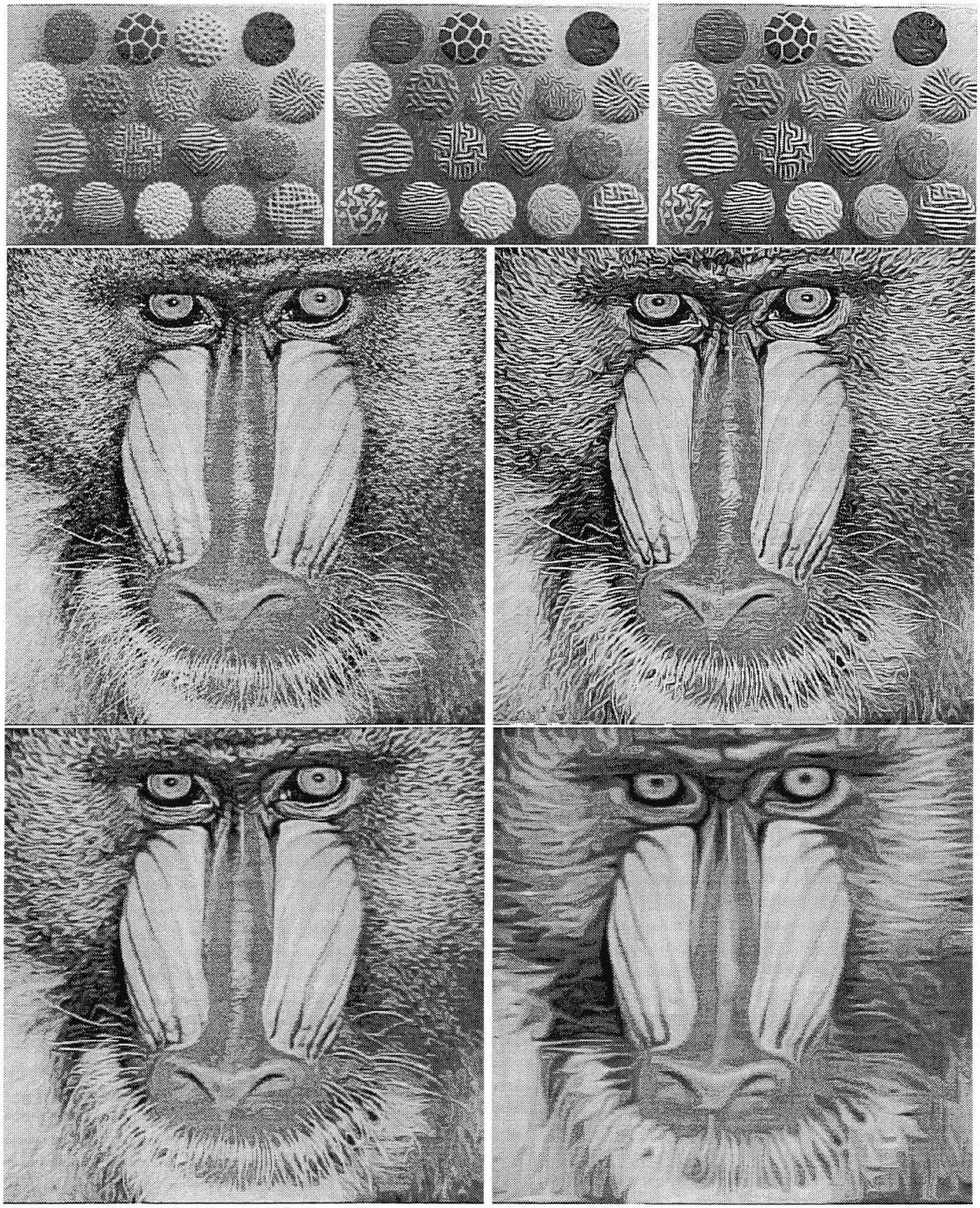

Figure 5: Color and texture reaction diffusion;Top Row: Original 'Shells' image 242×184 (left), and the result of the flow for 4 (middle) and 8 (right) numerical iterations, $\alpha = 0.55$. Second Row: Original 'mandrill' image 512×512 (left), and the result of orientation-preserving flow and negative eigenvalue (reaction) in gradient direction, $\alpha = 0.39$. Third Row: Two steps along an orientation-preserving diffusion flow.

The discretization of Eq. (25) is given by

$$W_{p,q,l,m} = (T^{wav}_{p,q,l,m} I) = a^{-m} \int\int dxdy I(x,y)\psi_{l\Delta\theta}(a^{-m}(x-p\Delta x), a^{-m}(y-q\Delta x)), \quad (29)$$

where Δx is the basic sampling interval, and the angles are given by $\Delta\theta = 2\pi l/L$, $l = 0, ..., L-1$, L being the total number of orientations; p, q and m are integers determining the position and scaling. Note that as m increases the sample intervals get larger forming a pyramidal structure. Eq. 29 can be interpreted as a projection onto a discrete set of basis functions, namely

$$W_{p,q,l,m} = \langle I, \psi_{p,q,l,m} \rangle. \quad (30)$$

The real number k determines the frequency bandwidth of the filters in octaves via the approximation

$$k = \frac{a^{\phi}+1}{a^{\phi}-1}\sqrt{2\ln 2}, \quad (31)$$

where ϕ is the bandwidth in octaves, e.g. for $a = 2$ and $\phi = 1.5$ we get $k \approx 2.5$. In the above approximation the DC normalization term $e^{-k^2/2}$ that is required to make a wavelet basis out of the Gabor basis is ignored and we consider $a = k/\omega_0$. So the peaks of the scaled mother wavelets in the frequency domain are (approximately) at the locations $a^{-m}\omega_0$.

For our application we have chosen $L = 16$ (16 orientations), $a = 2$, $\Delta x = 1$, $k = 2.5$, and 5 scales, i.e. $m \in \{0, .., 4\}$. This selection results in a 'tight frame' [11] that allows simple summation reconstruction.

5.1 Beltrami Flow for Texture Enhancement

We denote the 2D Gabor/Morlet-wavelet transform as $W(x, y, \theta, \sigma)$, where for the discrete case $\sigma = a^m$ and $\theta = l\Delta\theta$. Let $R = \text{Real}(W)$ and $J = \text{Imag}(W)$ be its real and imaginary part. The response of a simple cell is then modeled by the projection of the image onto a specific Gabor/Morlet wavelet.

The Gabor/Morlet-wavelet transform of an image in our framework is a mapping $W : (x, y, \theta, \sigma) \rightarrow (x, y, \theta, \sigma, R, J)$, i.e. a $4D$ manifold embedded in $6D$. The Beltrami operator is not limited to act on gray level images (2D surfaces in 3D) as we have shown in Section 3 for color. First, the metric $g_{\mu\nu}$ is "pulled back" from the relevant arclength definition in the spatial-orientation complex space, namely

$$ds^2 = dx^2 + dy^2 + d\theta^2 + d\sigma^2 + dJ^2 + dR^2.$$

For practical implementation we consider each scale as a separate space. This is in contrast to writing the arclength for the full transform. Therefore, the arclength for a given scale σ is $ds^2 = dx^2 + dy^2 + d\theta^2 + dJ^2 + dR^2$, and the induced metric for each scale is given by

$$(g_{\mu\nu}) = \begin{pmatrix} 1 + R_x^2 + J_x^2 & R_xR_y + J_xJ_y & R_xR_\theta + J_xJ_\theta \\ R_xR_y + J_xJ_y & 1 + R_y^2 + J_y^2 & R_yR_\theta + J_yJ_\theta \\ R_xR_\theta + J_xJ_\theta & R_yR_\theta + J_yJ_\theta & 1 + R_\theta^2 + J_\theta^2 \end{pmatrix}. \quad (32)$$

As we have seen before, the above result can be understood from the arclength definition and applying the chain rule $dR = R_x dx + R_y dy + R_\theta d\theta$, and similarly for dJ to obtain the desired bilinear structure.

Finally, the area-minimizing and feature-preserving *Beltrami flow* that operates on the Gabor/Morlet-wavelet transform of a texture image can be compactly written as

$$\begin{aligned} R_t &= \Delta_g R \\ J_t &= \Delta_g J. \end{aligned} \tag{33}$$

5.2 Experimental Results

As a by product of the wavelet decomposition, at each scale σ we now have the complex function $W_\sigma(x, y, \theta) = R_\sigma(x, y, \theta) + iJ_\sigma(x, y, \theta)$. It defines a 3D manifold in the 5D space $(x, y, \theta, R_\sigma, J_\sigma)$. The extra coordinate θ that describes the behavior of the image along a specific direction enables us to smooth the image while keeping the meaningful orientation structure of the texture. Moreover, we have the freedom to apply different filters to the different scales. This enables us to preserve the nature of texture images by processing them only at significant scales. In other words, we can sharpen a specific scale without effecting the rest of the sub-band images. The "scale" at which we choose to apply the filter is similar to the role of the parameter ρ, the variance of the Gaussian used to smooth the metric $g_{\mu\nu}$ in Section 4. The first row is Fig. 6 presents the original image and the result of applying the Beltrami flow in the decomposition space to filter out non-oriented structures in a gray level image. More examples are shown in the second and third rows of Fig. 6.

6 Movies and Volumetric Medical Images

Traditionally, MRI volumetric data is referred to as 3D medical image. Following our framework, a more appropriate definition is of a 3D surface in 4D (x, y, z, I). In a very similar manner we will consider gray level movies as a 3D surfaces in 4D, where all we need to do is the mental exercise of replacing z of the volumetric medical images by the sequence (time) axis. In Fig. 7, the first row shows images at different z locations and the second row shows the corresponding denoised images. This is a relatively simple case, since now we have co-dimension equal to one.

The induced metric in this case is given by

$$(g_{ij}) = \begin{pmatrix} 1 + I_x^2 & I_x I_y & I_x I_z \\ I_x I_y & 1 + I_y^2 & I_y I_z \\ I_x I_z & I_y I_z & 1 + I_z^2 \end{pmatrix}, \tag{34}$$

and the Beltrami flow is:

$$I_t = \frac{1}{\sqrt{g}} \mathrm{div} \left(\frac{\nabla I}{\sqrt{g}} \right), \tag{35}$$

where now $\nabla I \equiv (I_x, I_y, I_z)$ and $g = 1 + I_x^2 + I_y^2 + I_z^2$.

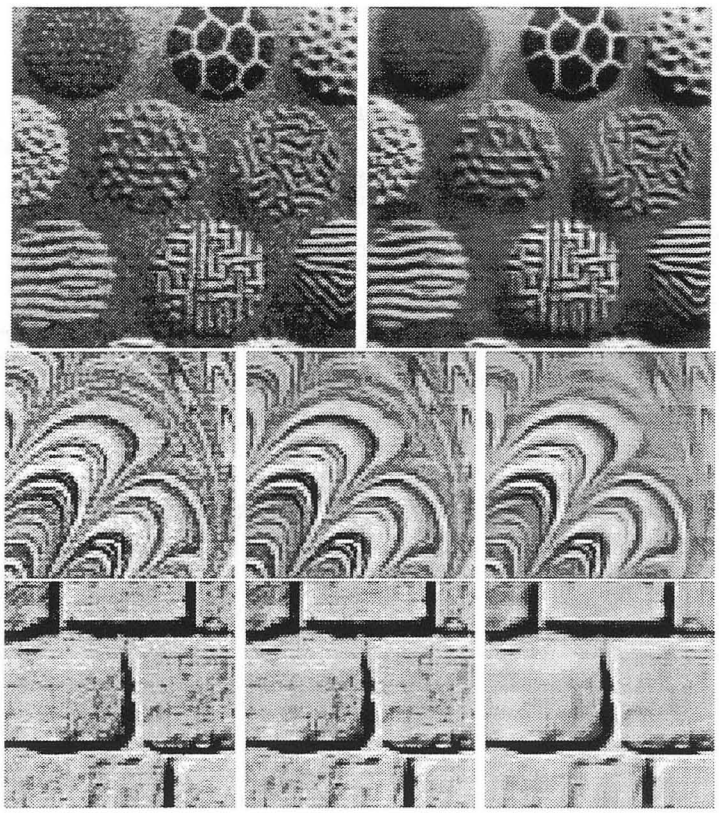

Figure 6: Top Row: Original image 128×128 is on the left, Result of Beltrami flow for 70 numerical iterations of each sub-scale in the decomposition space is on the right. Second and Third Rows: Two steps along the evolution for two different texture images, Left is the original image 64×64.

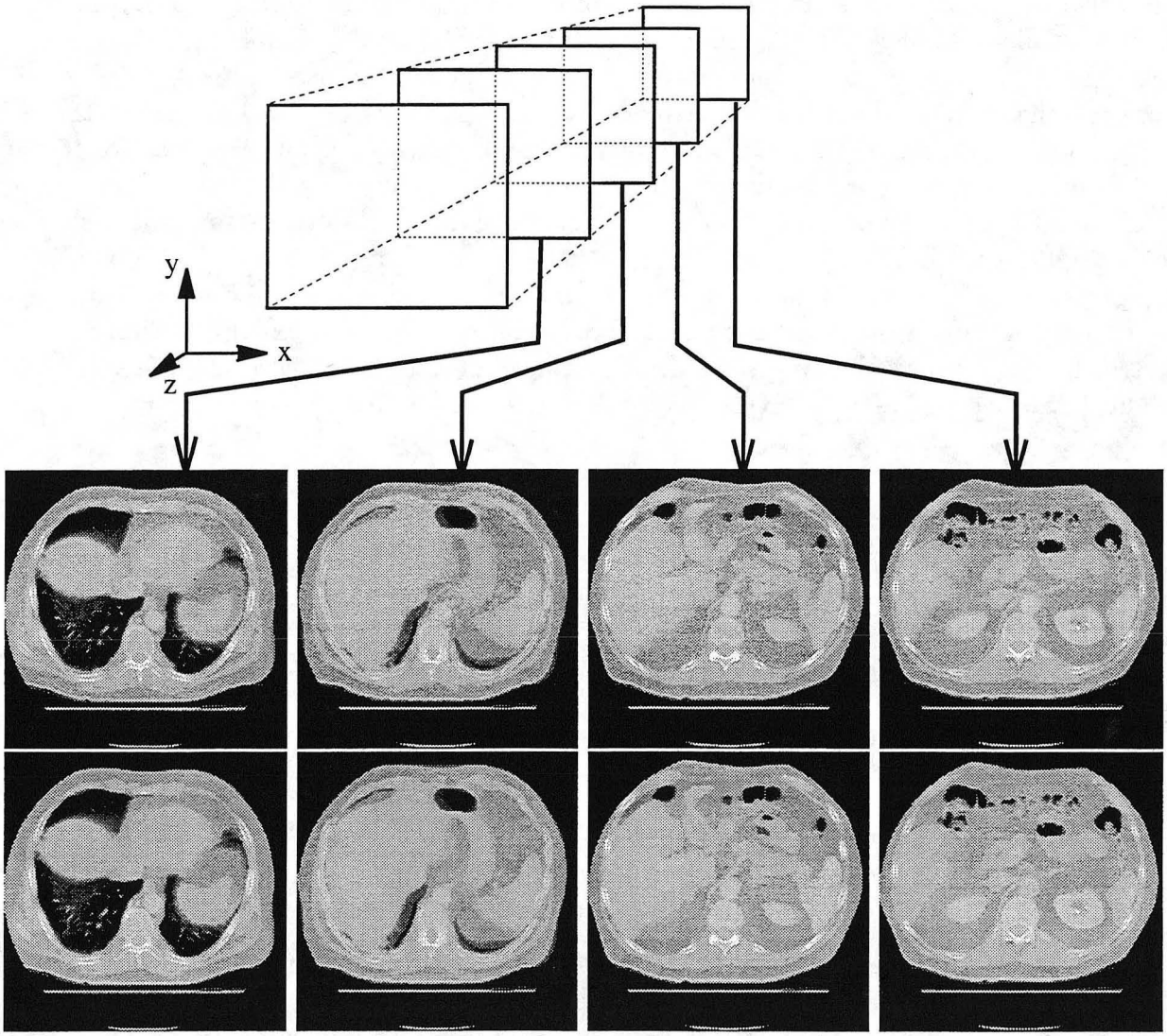

Figure 7: Movie or volumetric data; see text.

7 Concluding Remarks

We introduced a geometric framework and used it to design novel procedures for coherence enhancement of color and gray level images. These procedures are based on the interpretation of the image as a surface and a heat flow with respect to a given metric (Beltrami operator) as a filter.

We dealt with image enhancement and reconstruction of color and orientation based texture. These two different spaces were linked by a geometrical measure. The proposed filters align the color channels without un-coupling disturbances while enhancing the orientation based texture features and/or preserving the edges. In one of the examples Weickert [35] texture enhancement algorithm was extended to texture-color and was linked it to the geometric framework. In another example Lee's [17] decomposition space was used for texture processing, again, via the geometric framework.

A direct application of the proposed method is to enhance, selectively smooth, or sharpen color-texture and volumetric images. It can also be used to reduce the image entropy prior to compression and enhance its coherence in the reconstruction process (e.g. restoring images and denoising lossy compression effects). It was shown by several examples, that the geometrical framework can be applied to color, texture, multi channel data, movies, volumetric medical data, as well as non-trivial decomposition spaces.

References

[1] L Alvarez and L Mazora. Signal and image restoration using shock filters and anisotropic diffusion. *SIAM J. Numer. Anal*, 31:590-605, 1994.

[2] A Blake and A Zisserman. *Visual Reconstruction.* MIT Press, Cambridge, Massachusetts, 1987.

[3] P Blomgren and T F Chan. Color TV: Total variation methods for restoration of vector valued images. CAM TR, UCLA 1996.

[4] A C Bovik, M Clark, and W S Geisler. Multichannel texture analysis using localized spatial filters. *IEEE Trans. on PAMI*, 12(1):55–73, 1990.

[5] V Caselles, R Kimmel, and G Sapiro. Geodesic active contours. In *Proc. ICCV'95*, pages 694–699, Boston, Massachusetts, June 1995.

[6] T F Chan, G H Golub, and P Mulet. A nonlinear primal-dual method for total variation-based image restoration. Presented at AMS/SIAM workshop on Linear and Nonlinear CG methods, July 1995.

[7] A Chambolle. Partial differential equations and image processing. In *Proc. IEEE ICIP*, Austin, Texas, November 1994.

[8] G H Cottet and L Germain. *Image processing through reaction combined with nonlinear diffusion.* Math. Comp. Vol. 61, 659–673, 1993.

[9] I Daubechies. The wavelet transform, time frequency localization and signal analysis. *IEEE Trans. Information Theory*, 36(5):961–1004, 1990.

[10] S Di Zenzo. A note on the gradient of a multi image. *Computer Vision, Graphics, and Image Processing*, 33:116–125, 1986.

[11] R J Duffin and A C Schaeffer. A class of nonharmonic Fourier series. *Trans. Am. Math. Soc.*, 72:341–366, 1952.

[12] A I El-Fallah, G E Ford, V R Algazi, and R R Estes. The invariance of edges and corners under mean curvature diffusions of images. In *Processing III SPIE*, volume 2421, pages 2–14, 1994.

[13] D Gabor. Information theory in electron microscopy. *Laboratory Investigation*, 14(6):801–807, 1965.

[14] R Kimmel, What is a natural norm for multi channel image processing. Report LBNL-40097, UC-405, Lawrence Berkeley National Laboratory CA 94720, March 1997. submitted.

[15] R Kimmel, N Sochen, and R Malladi. Images as embedding maps and minimal surfaces: Movies, color, and volumetric medical images. In Proc. of *IEEE CVPR'97*, 1997.

[16] E Kreyszing. *Differential Geometry.* Dover Publications, Inc., New York, 1991.

[17] T S Lee. Image representation using 2D Gabor-wavelets. *IEEE Trans. on PAMI*, 18(10):959–971, 1996.

[18] T S Lee, D Mumford, and A L Yuille. Texture segmentation by minimizing vector valued energy functionals: the couple-membrane model. In G Sandini, editor, *Lecture Notes in Computer Science, 588, Computer Vision: ECCV'92*, pages 165–173. Springer-Verlag, 1992.

[19] M Lindenbaum, M Fischer, and A M Bruckstein. On Gabor's contribution to image enhancement. *Pattern Recognition*, 27(1):1–8, 1994.

[20] R Malladi and J A Sethian. Image processing: Flows under min/max curvature and mean curvature. *Graphical Models and Image Processing*, 58(2):127–141, March 1996.

[21] B S Manjunath and W Y Ma. Texture features for browsing and retrieval of image data. *IEEE Trans. on PAMI*, 18(8):837–841, 1996.

[22] T Masters. *Signal and image processing with neural networks: A C++ Sourcebook.* Wiley, New York, 1996.

[23] D Mumford and J Shah. Boundary detection by minimizing functionals. In *Proc. of CVPR*, San Francisco, 1985.

[24] S J Osher and L I Rudin. Feature–Oriented Image Enhancement Using Shock Filters. *SIAM J. Numer. Analy.*, 27(4):919-940, 1990.

[25] P Perona and J Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE-PAMI*, 12:629–639, 1990.

[26] A M Polyakov. *Physics Letters*, 103B:207, 1981.

[27] M Porat and Y Y Zeevi. The generalized Gabor scheme of image representation in biological and machine vision. *IEEE Trans. on PAMI*, 10(4):452–468, 1988.

[28] M Proesmans, E Pauwels, and L van Gool. Coupled geometry-driven diffusion equations for low level vision. In B M ter Haar Romeny, editor, *Geometric–Driven Diffusion in Computer Vision.* Kluwer Academic Publishers, The Netherlands, 1994.

[29] Y Rubner and C Tomasi. Coalescing texture descriptors. In *Proc. of the ARPA Image Understanding Workshop*, Feb. 1996.

[30] L Rudin, S Osher, and E Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[31] G Sapiro. Vector-valued active contours. In *Proc. IEEE CVPR'96*, pages 680–685, 1996.

[32] G Sapiro and D Ringach. Anisotropic diffusion of multivalued images. In *12th Int. Conf. on Analysis and Optimization of Systems: Images, Wavelets and PDE'S*, Paris, June 1996. Springer Verlag.

[33] G Sapiro and D L Ringach. Anisotropic diffusion in color space. *IEEE Trans. Image Proc.*, 5:1582-1586, 1996.

[34] N Sochen, R Kimmel, and R Malladi. From high energy physics to low level vision. Report LBNL 39243, Lawrence Berkeley National Laboratory, CA 94720, August 1996. Short version in press: *IEEE Tran. on IP* 1997.

[35] J Weickert. *Anisotropic diffusion in image processing.* Ph.D. thesis, Kaiserslautern Univ., Kaiserslautern, Germany, November 1995.

[36] J Weickert. *Multiscale texture enhancement.* In *Computer analysis of images and patterns; Lecture Notes in Computer Science*, Vol. 970, Springer, pp. 230-237, 1995.

[37] R Whitaker and G Gerig. Vector-valued diffusion. In B M ter Haar Romeny, editor, *Geometric–Driven Diffusion in Computer Vision.* Kluwer Academic Publishers, The Netherlands, 1994.

[38] S D Yanowitz and A M Bruckstein. A new method for image segmentation. *Computer Vision, Graphics, and Image Processing*, 46:82–95, 1989.

[39] A. Yezzi. Modified curvature motion for image smoothing and enhancement. *IEEE Trans. IP*, to appear, 1997.

ERNEST ORLANDO LAWRENCE BERKELEY NATIONAL LABORATORY
ONE CYCLOTRON ROAD | BERKELEY, CALIFORNIA 94720

Prepared for the U.S. Department of Energy under Contract No. DE-AC03-76SF00098