LBNL-39640
UC-405
Preprint

ERNEST ORLANDO LAWRENCE BERKELEY NATIONAL LABORATORY

On the Geometry of Texture

R. Kimmel, N. Sochen, and R. Malladi
Computer Sciences Directorate
Mathematics Department

November 1996
To be submitted
for publication

REFERENCE COPY
Does Not
Circulate
Bldg. 50 Library.
Copy 1
LBNL-39640

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-39640

ON THE GEOMETRY OF TEXTURE*

R. Kimmel
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720, USA

N. Sochen
Raymond and Beverly Sackler Faculty of Exact Sciences
Tel-Aviv University
Tel-Aviv, Israel

R. Malladi
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720, USA

November 1996

*This work was supported in part by the Applied Mathematical Sciences subprogram of the Office of Energy Research, U.S. Department of Energy, under Contract Number DE-AC03-76SF00098, and in part by the Office of Naval Research under Grant N00014-96-1-0381.

On the Geometry of Texture *

Ron Kimmel Nir Sochen Ravi Malladi

Abstract

We consider texture images as a composition of manifolds in the feature-space. This geometrical interpretation leads to a natural way for texture enhancement. A flow, based on manifold volume minimization yields a natural enhancement procedure for texture images. The 2D Gabor-Morlet transform is first used to decompose the image into sub-band images, where each sub-image corresponds to a different scale. Each sub-band image may be considered as a 3D manifold in a 5D space from which the original image can be reconstructed in a numerically stable way. Following our previous results, we then invoke Polyakov action from String theory, and develop a minimization process through a geometric flow that efficiently enhances each sub-band image in a spatial-orientation feature space. Finally, the enhanced sub-band images are composed back into an enhanced texture image.

1 Introduction

Texture plays an important role in the understanding process of many images. 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.

The psychophysical research of these first steps focuses on the way the brain cells are activated under the stimulus of a given image. Such experiments combined with recent developments in the field of signal representation led to relatively simple mathematical models that simulate the first steps in the way our brain represents images. One such model is based on the 2D Gabor/Morlet-wavelet transform of the image. Some nice mathematical properties and the relation of this transform to the physiological behavior were studied in [11, 17]. This model was used for the segmentation, interpretation and analysis of texture [2, 12], for texture based browsing [14], etc.

*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.

In this paper we use the same space to represent texture images. Then, we search for a geometrical way to improve and enhance texture based images. The geometrical feature enhancement procedure we introduce may serve as a step towards segmentation. This procedure is based on a flow in the transformed space in which the transform coefficients are treated as higher dimensional manifolds. A special minimization process, preserves domains of constant/homogeneous texture, enhances the texture in each domain, and thereby sharpens the boundaries between neighboring domains with different textures.

The remainder of this paper is organized as follows: Section 2 briefly reviews our previous results: the definition of arclength, the consideration of images as surfaces, and the minimization of Polyakov action that leads to a geometric flow we named the *Beltrami flow*. Next Section 3 describes the relevant feature space to the texture case. It gives the basics for constructing the 2D Gabor-Morlet wavelet decomposition, and a simple way for composing the image back. Section 4 presents experimental results of the Beltrami flow in the decomposition feature space, for simple gray level texture as well as color texture.

2 Images as Embedded Maps that flow toward Harmonic Maps

In [20] we consider images as 2D surfaces in higher dimensional spaces. We construct enhancement and segmentation procedures for color images as 2D surfaces in 5D $(x, y, \mathbf{r}, \mathbf{g}, \mathbf{b})$ space. As shown in [9] the idea of images as curved spaces is not limited to 2D surfaces, so that movies and volumetric images can be considered as 3D hyper surfaces (manifolds) in 4D $(x, y, z, I(x, y, z))$ space.

Our geometric framework finds a seamless link between the L_1 (Osher-Rudin TV and its variants) and the L_2 norms (used in Mumford-Shah 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 image plane, is the switch between the two commonly used norms. This observa-

[1] TV (Total Variation) schemes are based on minimizing the L_1 norm, namely $\int |\nabla I|$, while the L_2 norm minimizes $\int |\nabla I|^2$.

tion 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 yield very promising results for color image enhancement [20].

In this work, we 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 [18, 4, 19, 22]; see also [21] for orientation-preserving flows on the image itself. All these approaches begin with a flat metric [6] 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 perception 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. It will help us to construct a geometric flow in the 2D Gabor/Morlet wavelet transform that results in a high dimensional space. Other schemes that consider the image as a surface were proposed [1, 8, 23, 13], some even used the image information to build a Riemannian metric for segmentation [3]. However, non of these methods were generalized to feature space or any co-dimension higher than one.

2.1 The Metric

The basic concept of Riemannian differential geometry is distance. Let us start with the important example $X : \Sigma \rightarrow \mathbb{R}^3$. We denote the local coordinates on the two dimensional manifold Σ by (σ^1, σ^2). The map X is explicitly given by $(X^1(\sigma^1,\sigma^2), X^2(\sigma^1,\sigma^2), X^3(\sigma^1,\sigma^2))$. 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 $g_{ij}(\sigma^1,\sigma^2)$ called the metric

$$\begin{aligned} ds^2 &= g_{\mu\nu} d\sigma^\mu d\sigma^\nu \\ &\equiv g_{11}(d\sigma^1)^2 + 2g_{12}d\sigma^1 d\sigma^2 + g_{22}(d\sigma^2)^2, \end{aligned} \quad (1)$$

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 δ^μ_γ is the Kronecker delta.

[2] This flat metric is called 'structure tensor' in [21].

2.2 Induced metric

Let $X : \Sigma \rightarrow M$ be an embedding of Σ in M, where M is a Riemannian manifold with a metric $(g_{ij})_M$. We can use the knowledge of the metric on M and the map X to construct the metric on Σ. This procedure, is called the *pullback* and is given explicitly as follows:

$$(g_{\mu\nu})_\Sigma(\sigma^1,\sigma^2) = (g_{ij})_M(X(\sigma^1,\sigma^2))\partial_\mu X^i \partial_\nu X^j, \quad (2)$$

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

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 $\mathbb{R}^3$,

$$X : (\sigma^1,\sigma^2) \rightarrow (\sigma^1,\sigma^2, I(\sigma^1,\sigma^2)). \quad (3)$$

Using Eq. (2) we get

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

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

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

2.3 Polyakov Action

Let us briefly review our general framework for nonlinear diffusion in computer vision. We will use this framework in Section 3 to diffuse a textured image in the transformed domain. The equations will be 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 $X : \Sigma \rightarrow M$ has the following weight

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

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. This functional, for $m = 2$, was first

proposed by Polyakov [16] in the context of high energy physics, and the theory known as *string theory*.

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 [20] 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 variation calculus (see [20]), the Euler-Lagrange equations with respect to the embedding are:

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

Few remarks are in order. First 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 doesn't 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* [10], or for short *Beltrami operator*, and we will denote it 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

$$g_{\mu\nu} = \partial_\mu X^i \partial_\nu X_i. \qquad (7)$$

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

$$\vec{X}_t = H\vec{\mathcal{N}}, \qquad (8)$$

where H is the mean curvature, $\vec{\mathcal{N}}$ is the normal to the surface: [3]

$$H = \frac{(1+I_x^2)I_{yy} - 2I_xI_yI_{xy} + (1+I_y^2)I_{xx}}{g^{\frac{3}{2}}},$$

$$\vec{\mathcal{N}} = \frac{1}{\sqrt{g}}(-I_y, -I_x, 1)^T, \qquad (9)$$

and $g = 1 + I_x^2 + I_y^2$. We see that this choice gives us the mean curvature flow! 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 $X : \Sigma \to M$ 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 using the induced metric Eq. 7; see [20] for more details.

2.4 The Beltrami flow

Let us present a new and natural flow for images as surfaces. First let consider the case in which the gray level image is regarded as an embedding map $X : \Sigma \to \mathbb{R}^3$, where Σ is a two dimensional manifold, and the flow is natural in the sense that it minimizes the action functional with respect to I and (g_{ij}), 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:

$$I_t = \Delta_g I \equiv \frac{1}{\sqrt{g}}\partial_\mu(\sqrt{g}g^{\mu\nu}\partial_\nu I) = H\vec{\mathcal{N}}_{\hat{I}} \qquad (10)$$

where the metric is the induced one given in Eq. 4, and $\hat{I}$ is the unit vector in the I direction.

The geometrical meaning is obvious. Each point on the image surface moves with a velocity that depends on the mean curvature and the I component of the normal to the surface 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 image towards a minimal surface at a more rapid rate.

The Beltrami operator is not limited to act on gray level images (2D surfaces in 3D). In what follows we apply this operator to construct a feature preserving flow, e.g. the *Beltrami flow* of the coupled maps

[3] Note also that some definitions of the mean curvature include a factor of 2 that we omit in our definition.

$A(x,y,z)$ and $B(x,y,z)$ that define a manifold in the 5D space (x,y,z,A,B), is given by

$$\begin{aligned} A_t &= \nabla_g A \\ B_t &= \nabla_g B. \end{aligned} \tag{11}$$

The metric g is 'pulled back' from the relevant arclength definition, e.g.

$$ds^2 = dx^2 + dy^2 + dz^2 + dA^2 + dB^2.$$

3 2D Gabor/Morlet-wavelets: A Natural Space for Texture Images

In [11] Lee argues that the 2D Gabor/Morlet wavelet transform with specific coefficients is an appropriate mathematical description for images. He based his findings on recent neurophysiological evidence based on experiments on the visual cortex of mammalian brains. These experiments indicate that the best model for the filter response of simple cells are self-similar 2D Gabor/Morlet wavelets.

A full review of non-orthogonal wavelets, frames and frame bounds, is beyond the scope of this paper. We refer the interested reader to [15] for implementation considerations, and to the rich literature on wavelet theory, e.g. [5]. Here, we will comment on the basic concepts that are relevant to our discussion.

Following Lee [11], let us briefly describe the 2D Gabor/Morlet wavelets that model the simple cells while satisfying Daubechies' wavelet theory [5]. The 2D wavelet transform on an image $I(x,y)$, is defined as

$$(T^{wav}I)(x_0,y_0,\theta,a) =$$

$$\|a\|^{-1}\int\int dxdyI(x,y)\psi_\theta\left(\frac{x-x_0}{a},\frac{y-y_0}{a}\right), \tag{12}$$

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

$$\psi(x,y,x_0,y_0,\theta,a) = \|a\|^{-1}\psi_\theta(\frac{x-x_0}{a},\frac{y-y_0}{a}), \tag{13}$$

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

$$\psi(x,y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{8}(4x^2+y^2)}(e^{ikx} - e^{-\frac{k^2}{2}}), \tag{14}$$

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

$$\begin{cases} \tilde{x} &= x\cos\theta + y\sin\theta \\ \tilde{y} &= -x\sin\theta + y\cos\theta. \end{cases} \tag{15}$$

The discretization of Eq. (12) is given by

$$W_{p,q,l,m} = (T^{wav}_{p,q,l,m}I) =$$

$$a^{-m}\int\int dxdyI(x,y)\psi_{l\Delta\theta}(a^{-m}(x-p\Delta x), a^{-m}(y-q\Delta x)), \tag{16}$$

where Δx is the basic sampling interval, and the angles are given by $\Delta\theta = 2\pi l/L$, where $l = 0,...,L-1$ and L is 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. 16 can be read as a projection onto a discrete set of basis functions

$$W_{p,q,l,m} = \langle I, \psi_{p,q,l,m}\rangle. \tag{17}$$

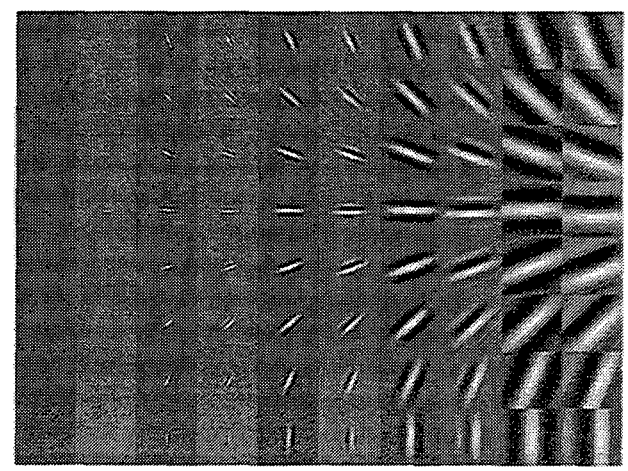

Figure 1: The wavelet basis functions (up to translations). The basis functions are presented in a gray level array, real (symmetric) and imaginary (asymmetric) for the 8 angles $[0,\pi]$ and 5 scales.

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

$$k = \frac{a^\phi+1}{a^\phi-1}\sqrt{2\ln 2}, \tag{18}$$

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,

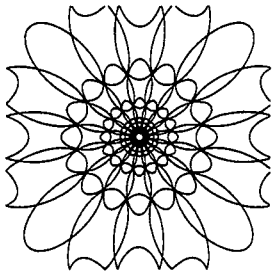

Figure 2: The half peak contours in the frequency domain of the wavelet basis functions in the previous Figure, (5 scales 16 orientations).

i.e. $m \in \{0,..,4\}$. This selection results in a frame bounds ratio of $B/A = 1.19$. The fact that this ratio is close to 1 means that we have a tight frame that allows simple summation reconstruction. Figs. 1 and 2 show the basis functions we used.

The concept of frames was introduced in [7]. It states that a family of functions (ψ_j) is a *frame* if there exist $A > 0, B < \infty$ that are called *frame bounds* so that for every f we have

$$A\|f\|^2 \leq \sum_j |\langle f, \psi_j \rangle|^2 \leq B\|f\|^2,$$

where $\|f\| = \int f^2$. One could recognize it as a generalization of Parseval's theorem. A discrete family of wavelets that forms a frame provides a complete representation of any function. In some cases it is possible to recover a function with good approximation by the inversion formula

$$f \approx \frac{2}{A+B} \sum_j \langle \psi_j, f \rangle \psi_j. \tag{19}$$

The ratio B/A measures the tightness of the frame. When $A = B$, the frame is tight and the reconstruction by summation is exact. Thus, as B/A approaches 1 we may still use the above reconstruction equation as a good approximation. That is, we treat our discrete wavelets as an orthonormal basis. For further analysis of the 2D Gabor/Morlet-wavelet transform, motivated by its close relation to the neurophysiological behavior of simple cells, we refer the reader to [11]. It gives the 2D extension of Daubechies [5] numerical analysis for some relevant frames and their bounds.

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

Motivated by the arrangement of simple cells in our brain, with as tight a frame as possible, we consider 5 spatial frequency octaves, and 16 angles that discretize the $[0, 2\pi]$ angular interval. Practically, we used the symmetry properties of the 2D Gabor/Morlet-wavelet transform: $W(x, y, \theta + \pi, \sigma) = \bar{W}(x, y, \theta, \sigma)$. Thus, only 8 angles are needed to represent the discretization of the full $[0, 2\pi]$ angular interval into 16. Periodic boundary conditions are used for the real (symmetric) part, and negative periodicity for the imaginary part, forming a 'Klein bottle' coordinate system in (x, y, θ). This enables us to reduce the memory complexity by a factor of 2.

The induced metric for the full transform; 4D manifold embedded in 6D space is given by

$$(g_{ij}) = \begin{pmatrix} 1+R_x^2+J_x^2 & R_xR_y+J_xJ_y & R_xR_\theta+J_xJ_\theta & R_xR_\tau+J_xJ_\tau \\ R_xR_y+J_xJ_y & 1+R_y^2+J_y^2 & R_yR_\theta+J_yJ_\theta & R_yR_\tau+J_yJ_\tau \\ R_xR_\theta+J_xJ_\theta & R_yR_\theta+J_yJ_\theta & 1+R_\theta^2+J_\theta^2 & R_\theta R_\tau+J_\theta J_\tau \\ R_xR_\tau+J_xJ_\tau & R_yR_\tau+J_yJ_\tau & R_\theta R_\tau+J_\theta J_\tau & 1+R_\tau^2+J_\tau^2 \end{pmatrix}, \tag{20}$$

where the space coordinates are (x, y, θ, τ, W) and $\tau \equiv \log\sigma$. For practical implementation that avoids the special numerical treatment needed along the pyramidal discrete σ scale axis, we consider each scale as a separate space. The induced metric for each scale is then given by

$$(g_{ij}) = \begin{pmatrix} 1+R_x^2+J_x^2 & R_xR_y+J_xJ_y & R_xR_\theta+J_xJ_\theta \\ R_xR_y+J_xJ_y & 1+R_y^2+J_y^2 & R_yR_\theta+J_yJ_\theta \\ R_xR_\theta+J_xJ_\theta & R_yR_\theta+J_yJ_\theta & 1+R_\theta^2+J_\theta^2 \end{pmatrix}. \tag{21}$$

This result can be understood from the arclength definition in this spatial-orientation complex space, namely

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

Applying the chain rule on $dR = R_x dx + R_y dy + R_\theta d\theta$, and similarly for dJ, we obtain the desired bilinear structure that describes the above induced metric for this case.

4 Experimental Results

Let us start with a simple example. In Figure 3 we first decompose an image via the wavelet transform into 4 separate sub-scale channels. The decomposition and the result of applying the Beltrami flow on each sub scale are shown.

Let us gain more motivation on the advantage of the wavelet decomposition. Figure 4 shows the result of composing the image back from just the first

2, and then the first 3 sub-scale channels. The cancelation of the shadowing can also be realized by a very simple high pass filter. However, as a by product of the wavelet decomposition, at each scale σ we now have the complex function $W_\sigma(x, y, \theta)$. It defines a surface in the 5D space (3 real and one complex dimensions) (x, y, θ, W_σ). 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 are able to sharpen a specific scale without effecting the rest of of the sub-band images. Fig. 5 is the original image and the result of applying the Beltrami flow to filter out non-oriented structures. More examples are shown in Fig. 6.

In the last example we deal with texture in color images. We first decompose each of the three (r, g, b) channels into its wavelet feature space as before. Denote the transformed complex functions as (W^r, W^g, W^b). Now we have 6 functions (a complex function for each channel) that map the (x, y, θ, σ) coordinates into a 10D feature space (4 coordinates plus 3 complex functions). In the following experiment we treat each scale separately, thus we deal with maps from 3D to 9D. The metric in this case is derived from the arclength

$$\begin{aligned} ds^2 &= dx^2 + dy^2 + d\theta^2 + dR^{r2} + dJ^{r2} + \\ &\quad dR^{g2} + dJ^{g2} + dR^{b2} + dJ^{b2}, \end{aligned} \tag{22}$$

where r, g, b stand for the three color channels, and R and J are the real (e.g. $R^r = \text{Real}(W^r)$) and imaginary (e.g. $J^r = \text{Imag}(W^r)$) parts of the decomposition for each channel. The metric for this case is given by its elements

$$\begin{aligned} g_{ij} &= \delta_{ij} + R^r{}_i R^r{}_j + J^r{}_i J^r{}_j + R^g{}_i R^g{}_j + \\ &\quad J^g{}_i J^g{}_j + R^r{}_i R^b{}_j + J^b{}_i J^b{}_j, \end{aligned} \tag{23}$$

where δ_{ij} is the Kronecker delta, and $R^r{}_i \equiv dR^r/di$ where $i, j \in \{x, y, \theta\}$. Fig. 7 is a color image and the result of the flow. Fig. 8 shows the effect on the three color channels. Fig. 9 is the flow effect on a larger color-texture image. This example demonstrates the natural way the proposed framework applies to even richer feature spaces.

5 Concluding Remarks

We proposed to combine a psychophysically supported texture space (via 2D Gabor/Morlet-wavelet

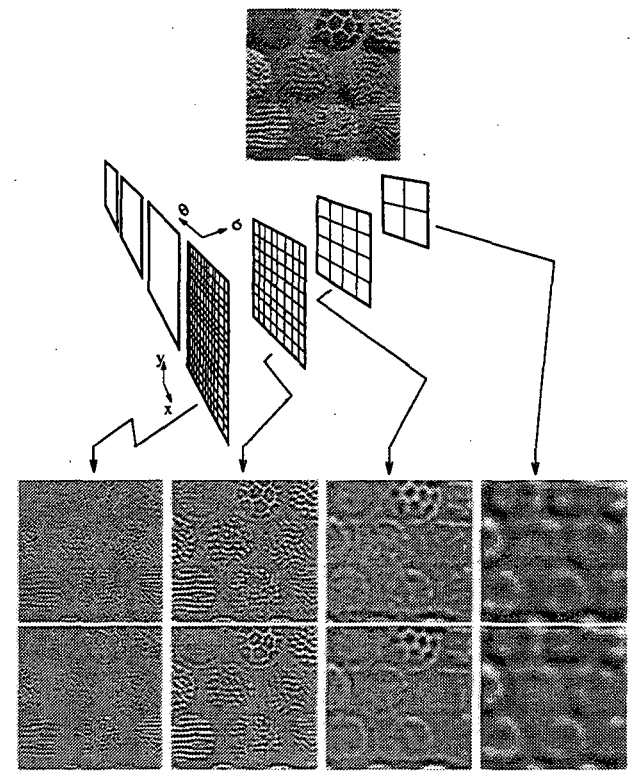

Figure 3: A schematic diagram of Gabor/Morlet wavelet decomposition of the original image (at the top) into the $(x, y, \theta, W_\sigma(x, y, \theta))$ and the images that are the result of reconstruction by summation for each scale σ separately (bottom). The last row presents the reconstruction result after 70 iteration of the Beltrami flow at each scale.

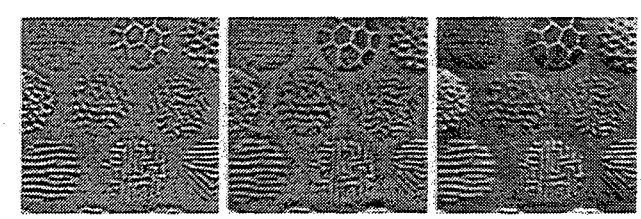

Figure 4: Reconstruction by summation, of only 2, 3, and all layers of the different scales: the low frequency scale contribute the shadowing, thus summing only over the first 3 scales cancels this effect (a simple high pass effect).

transform), with a geometrical flow to enhance texture images. The texture was considered as a manifold in its natural space. The flow was realized by invoking Polyakov action and the result was the Beltrami flow in the feature space. In our future work we intend to overcome the numerical complexity along the pyramidal logarithmic scale axis, and apply the flow in the full 6D $(x, y, \theta, \sigma, R, J)$ space (and the full 10D space for the color-texture space).

References

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

[2] 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.

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

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

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

[6] S Di Zenzo. A note on the gradient of a multi image. *CVGIP*, 33:116–125, 1986.

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

[8] 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 *Proc. III SPIE*, volume 2421, pages 2–14, 1994.

[9] R Kimmel, N Sochen, and R Malladi. Images as embedding maps and minimal surfaces: Movies, color, and volumetric medical images. *submitted to IEEE CVPR'97.*

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

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

[12] 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: ECCV'92*, pages 165–173. Springer-Verlag, 1992.

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

[14] 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.

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

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

[17] 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.

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

[19] G Sapiro and D L Ringach. Anisotropic diffusion in color space. *IEEE Trans. Image Proc.*, to appear, 1996.

[20] N Sochen, R Kimmel, and R Malladi. From high energy physics to low level vision. Report LBNL 39243, LBNL, UC Berkeley, CA 94720, Aug. 1996.

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

[22] 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.

[23] S D Yanowitz and A M Bruckstein. A new method for image segmentation. *CVGIP*, 46:82–95, 1989.

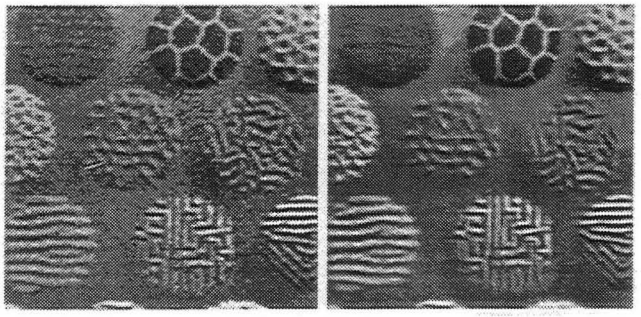

Figure 5: Left: Original image 128×128, Right: Result of Beltrami flow for 70 numerical iterations in each sub-scale.

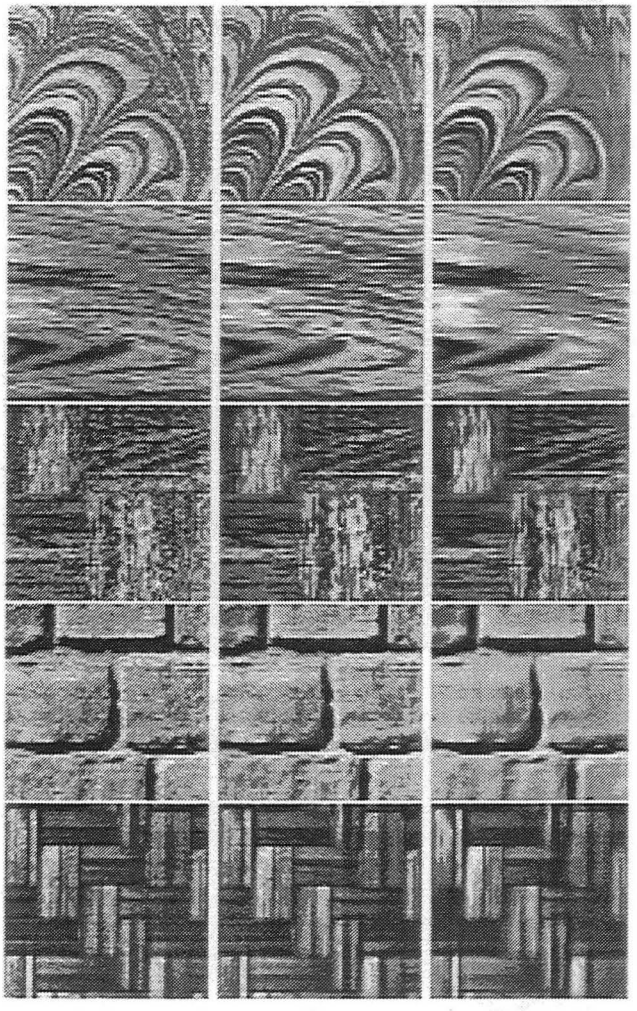

Figure 6: Example of 2 steps along the evolution for different texture images, Left: Original image 64×64.

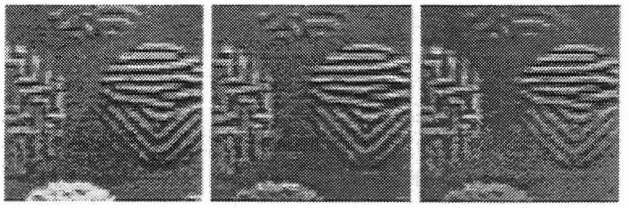

Figure 7: Texture and color: 2 steps along the evolution for a color texture image. The left is the original 64×64 image. *This is a color image.*

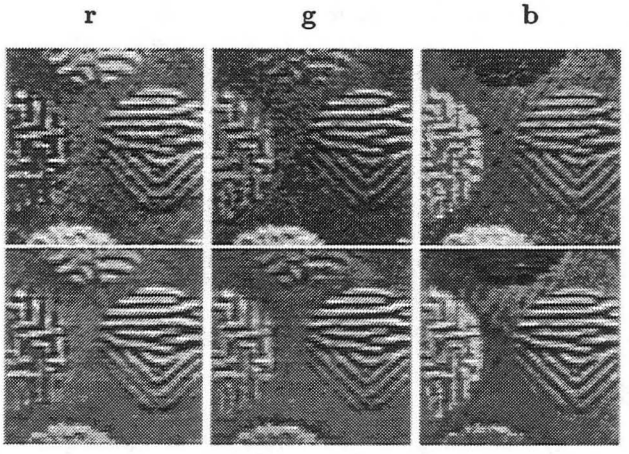

Figure 8: Texture and color; The three color channels of the flow result in the previous figure. Upper row: The original r, g, b channels (of the left most image in the previous figure). Bottom row: The r, g, b channels after 100 numerical iterations, (of the right image in the previous figure).

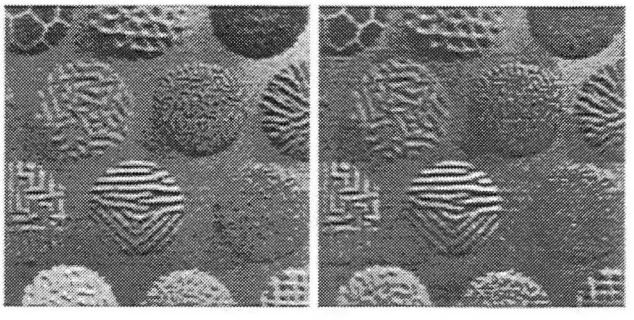

Figure 9: Texture and color; Right: original 128×128 color image. Left: The result after 70 numerical iterations. *This is a color image.*

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