On the Geometry of Texture

Ron Kimmel, Nir A. Sochen, and 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 [6,10]. This model was used for the segmentation, interpretation and analysis of texture [2,7], for texture based browsing [8], etc.

Curve and Surface Design: Saint-Malo 1999
Pierre-Jean Laurent, Paul Sablonnière, and Larry L. Schumaker (eds.), pp. 203–212.

ISBN 0-8265-1356-5.

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.

§2. Images as Embedded Maps that Flow Toward Harmonic Maps

In [11] 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 [4], the idea of images as curved spaces is not limited to 2D surfaces, so that movies and volumetric images can be considered as 3D hypersurfaces (manifolds) in 4D $(x, y, z, I(x, y, z))$ space.

Our geometric framework finds a seamless link between the L_1 norm, used in the Osher-Rudin TV image enhancement and its variants, and the L_2 norms, used in Mumford-Shah image segmentation and its variants. TV (Total Variation) schemes are based on minimizing the L_1 norm, namely $\int |\nabla I|$, while the L_2 norm minimizes $\int |\nabla I|^2$. Our framework is based on the geometry of the image and its interpretation as a surface. The aspect ratio between the gray level and the xy image plane, is the 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 yield very promising results for color image enhancement [11]. In this work, we propose a flow in a rich feature space which is different from the image spatial-intensity space.

Representation and Riemannian structure

We represent an image and other local features as embedding maps of a Riemannian manifold in a higher dimensional space. The simplest example is the image itself which is represented as a 2D surface embedded in $\mathbb{R}^3$. We denote the map by $X : \Sigma \to \mathbb{R}^3$, where Σ is a two-dimensional surface, and we denote the local coordinates on it by (σ^1, σ^2). The map X is given in general by $(X^1(\sigma^1, \sigma^2), X^2(\sigma^1, \sigma^2), X^3(\sigma^1, \sigma^2))$. In our example we represent it as

$(X^1 = \sigma^1, X^2 = \sigma^2, X^3 = I(\sigma^1, \sigma^2))$. We choose a Riemannian structure on this surface, namely, a metric. The metric is a positive definite and symmetric 2-tensor that may be defined through the local distance measurements:

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

where we used Einstein summation convention in the second equality. We denote the inverse of the metric by $g^{\mu\nu}$.

Polyakov action: a measure on the space of embedding maps

Let us briefly review our general framework for non-linear diffusion in computer vision. We will use this framework in Section 4 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 functional $S[X]$ attaches a real number to a map $X : \Sigma \to M$:

$$S[X^i, g_{\mu\nu}, h_{ij}] = \int dV \langle \nabla X^i, \nabla X^j \rangle_g h_{ij}, \tag{2}$$

where dV is a volume element and $\langle \nabla R, \nabla G \rangle_g = g^{\mu\nu} \partial_\mu R \partial_\nu G$. This functional, for $m = 2$, was first proposed by Polyakov [9] in the context of high energy physics, and the theory is known as string theory.

Using standard methods in the calculus of variations (see [11]), the Euler-Lagrange equations with respect to the embedding are

$$-\frac{1}{2\sqrt{g}} h^{il} \frac{\delta S}{\delta X^l} = \frac{1}{\sqrt{g}} \partial_\mu (\sqrt{g} g^{\mu\nu} \partial_\nu X^i). \tag{3}$$

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 reparameterization-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 [5], 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 this 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

$$g_{\mu\nu} = \partial_\mu X^i \partial_\nu X_i. \tag{4}$$

Plugging this induced metric in the first Euler-Lagrange equation (3), we get the steepest decent flow

$$\mathbf{X}_t = H\mathbf{N}, \tag{5}$$

where H is the mean curvature, and $\mathbf{N}$ is the normal to the surface given by

$$\begin{aligned} H &= \frac{(1+I_x^2)I_{yy} - 2I_xI_yI_{xy} + (1+I_y^2)I_{xx}}{g^{\frac{3}{2}}}, \\ \mathbf{N} &= \frac{1}{\sqrt{g}}(-I_y, -I_x, 1)^T, \end{aligned} \tag{6}$$

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

$$S = \int d\text{Area} = \int d^2\sigma\sqrt{g} = \int d^2\sigma\sqrt{\det(\partial_\mu X^i \partial_\nu X_i)}, \tag{7}$$

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.

§3. Gabor/Morlet-wavelets: A Natural Space for Texture Images

In [6] Lee argues that the 2D Gabor/Morlet wavelet transform with specific coefficients is an appropriate mathematical description for images. He based his findings on 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.

Following Lee [6], let us briefly describe the 2D Gabor/Morlet wavelets that model the simple cells. The 2D wavelet transform on an image $I(x, y)$ is defined as

$$(T^{wav}I)(x_0, y_0, \theta, a) = \|a\|^{-1} \int\int dxdyI(x,y)\psi_\theta\left(\frac{x-x_0}{a}, \frac{y-y_0}{a}\right), \tag{8}$$

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

$$\psi(x, y, x_0, y_0, \theta, a) = \|a\|^{-1}\psi_\theta(\frac{x-x_0}{a}, \frac{y-y_0}{a}), \tag{9}$$

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

$$\psi(x, y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{8}(4x^2+y^2)}(e^{ikx} - e^{-\frac{k^2}{2}}), \tag{10}$$

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

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}. \tag{11}$$

The discretization of (8), i.e $(T^{wav}_{p,q,l,m}I)$, is denoted by $W_{p,q,l,m}$ and is given by

$$W_{p,q,l,m} = a^{-m} \int\int dxdy I(x,y)\psi_{l\Delta\theta}(a^{-m}(x-p\Delta x), a^{-m}(y-q\Delta x)), \tag{12}$$

where Δx is the basic sampling interval, $\Delta\theta = 2\pi/L$, and the angles are given by $l\Delta\theta$, 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. Equation (12) can be read as a projection onto a discrete set of basis functions

$$W_{p,q,l,m} = \langle I, \psi_{p,q,l,m}\rangle. \tag{13}$$

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

$$k = \frac{a^\phi + 1}{a^\phi - 1}\sqrt{2\ln 2}, \tag{14}$$

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 to work with a frame. The concept of frames was introduced in [3]. 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

$$A\|f\|^2 \leq \sum_j |\langle f, \psi_j\rangle|^2 \leq B\|f\|^2, \tag{15}$$

where $\|f\| = \int f^2$. One could recognize this 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

$$f \approx \frac{2}{A+B}\sum_j \langle \psi_j, f\rangle \psi_j. \tag{16}$$

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

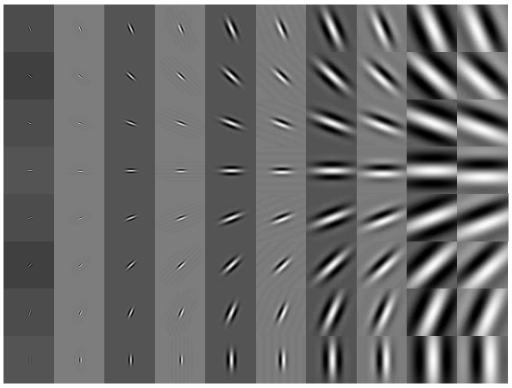

Fig. 1. The wavelet basis functions (up to translations). The basis functions are presented in a gray level array, real (symmetric) and imaginary (a-symmetric) for the 8 angles $[0, \pi]$ and 5 scales.

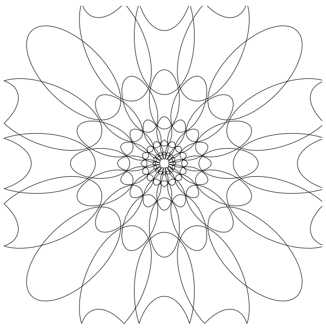

Fig. 2. The half peak contours in the frequency domain of the wavelet basis functions in the previous figure, (5 scales 16 orientations).

approximation. That is, we treat our discrete wavelets as an orthonormal basis.

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. We choose $a = 2$ and $\Delta x = 1$. This selection results in a frame bounds $A = 271.95$, $B = 233.69$, with 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. 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.

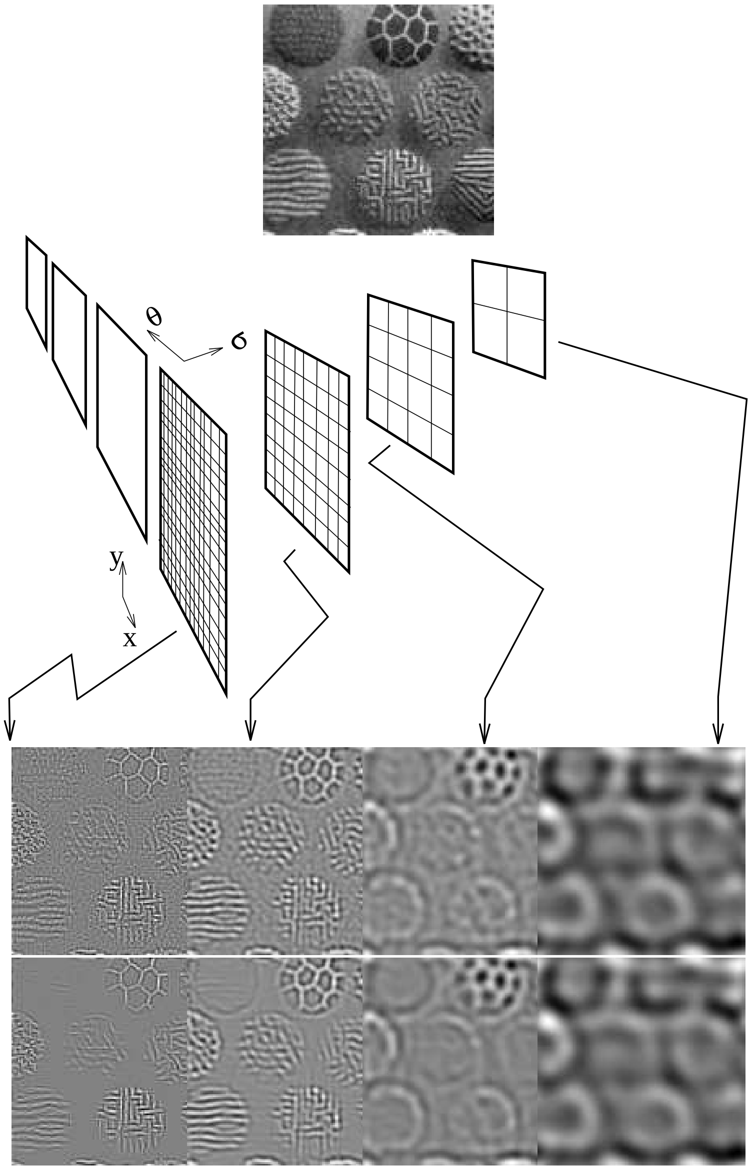

Fig. 3. A schematic diagram of Gabor/Morlet wavelet decomposition of the original image (at the top) into the $(x, y, \theta, W_\sigma(x, y, \theta))$ and the images that are the result of reconstruction by summation for each scale σ separately (bottom). The last row presents the reconstruction result after 70 iteration of the Beltrami flow at each scale. In all the examples we use $L = 16$, $a = 2$, $k = 2.5$, and $m \in \{0, .., 4\}$.

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

$$(g_{\mu\nu}) = \begin{pmatrix} 1 + R_x^2 + J_x^2 & R_x R_y + J_x J_y & R_x R_\theta + J_x J_\theta \\ R_x R_y + J_x J_y & 1 + R_y^2 + J_y^2 & R_y R_\theta + J_y J_\theta \\ R_x R_\theta + J_x J_\theta & R_y R_\theta + J_y J_\theta & 1 + R_\theta^2 + J_\theta^2 \end{pmatrix}. \qquad (17)$$

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

$$ds^2 = dx^2 + dy^2 + d\theta^2 + dJ^2 + dR^2. \tag{18}$$

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.

The gradient descent equations for the Polyakov action read

$$R_t = \Delta_g R, \qquad J_t = \Delta_g J, \tag{19}$$

where $\Delta_g X$ is given in (3) with the metric (17).

§4. Experimental Results

Let us start with a simple example. In Fig. 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. Fig. 4 shows the result of composing the image back from just the first 2, and then the first 3 sub-scale channels. The cancellation of the shadowing can also be realized by a very simple high pass filter. However, as a byproduct 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 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.

§5. Concluding Remarks

We proposed to combine a psychophysically supported texture space, the 2D Gabor/Morlet-wavelet 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. The result is a variational-geometric technique that enhances texture images in their appropriate decomposition space.

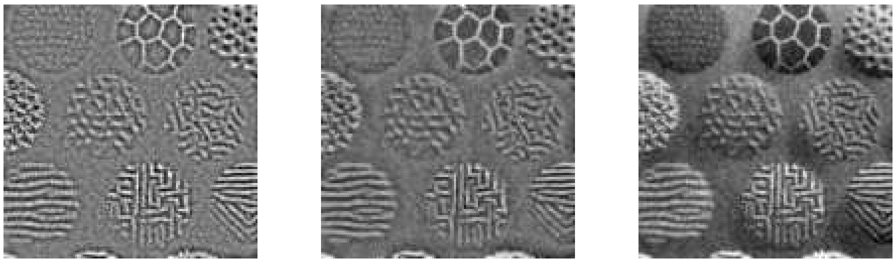

Fig. 4. Reconstruction by summation, of only 2, 3, and all layers of the different scales: the low frequency scale contribute the shadowing, thus summing only over the first 3 scales cancels this effect (a simple high pass effect).

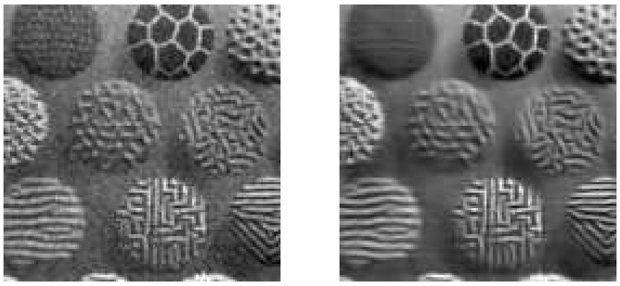

Fig. 5. Left: Original image 128×128, Right: Result of Beltrami flow for 70 numerical iterations in each sub-scale.

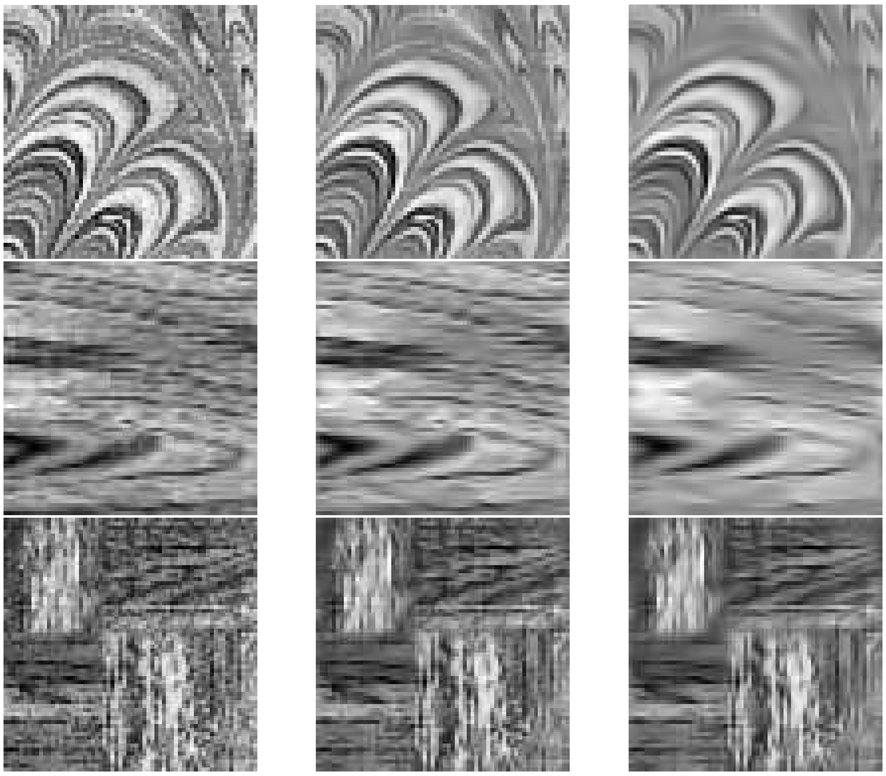

Fig. 6. Example of 2 snapshots from the evolution for different texture images Left: Original image 64×64.

References

1. Antoine, J. P., P. Carrette, R. Murenzi and B. Piette, Image analysis with 2-D continuous wavelet transform, Signal Processing **31**(3) (1993), 241–272.
2. Bovik, A. C., M. Clark, and W. S. Geisler, Multichannel texture analysis using localized spatial filters, IEEE Trans. Pattern Anal. and Machine Intelligence **12**(1) (1990), 55–73.
3. Duffin, R. J. and A. C. Schaeffer, A class of nonharmonic Fourier series, Trans. Am. Math. Soc. **72** (1952), 341–366.
4. Kimmel, R., N. Sochen, and R. Malladi, Images as embedding maps and minimal surfaces: Movies, color, and volumetric medical images, *IEEE CVPR'97*, Puerto Rico, June, 1997, 350–355.
5. Kreyszing, E., *Differential Geometry*, Dover Publications, Inc., New York, 1991.
6. Lee, T. S., Image representation using 2D Gabor-wavelets, IEEE Trans. on PAMI **18**(10) (1996), 959–971.
7. Lee, T., D. Mumford, and A. L. Yuille, Texture segmentation by minimizing vector valued energy functionals: the couple-membrane model, in Lecture Notes in Computer Science, 588: ECCV'92, G Sandini (ed.), Springer-Verlag, 1992, 165–173.
8. Manjunath, B. S., and W. Y. Ma, Texture features for browsing and retrieval of image data, IEEE Trans. Pattern Anal. and Machine Intelligence **18**(8) (1996), 837–841.
9. Polyakov, A. M., Physics Letters **103B** (1981), 207.
10. Porat, M. and Y. Y. Zeevi, The generalized Gabor scheme of image representation in biological and machine vision, IEEE Trans. on PAMI **10**(4) (1988), 452–468.
11. Sochen, N., R. Kimmel, and R. Malladi, From high energy physics to low level vision, Report LBNL 39243, LBNL, UC Berkeley, CA 94720, Aug. 1996.

Ron Kimmel, Computer Science Department
Technion, Israel Institute of Technology, Haifa 32000, Israel
`ron@cs.technion.ac.il`

Nir A. Sochen, School of Mathematical Sciences
Tel-Aviv University, Tel-Aviv 69978, Israel
`sochen@math.tau.ac.il`

Ravi Malladi, Lawrence Berkeley National Laboratory
University of California, Berkeley, CA 94720, USA
`malladi@math.lbl.gov`