Technical University of Denmark

DTU

Geometric Total Variation for Texture Deformation

Bespalov, Dmitriy; Dahl, Anders Bjorholm; Shokoufandeh, Ali

Published in:
Proceedings of 20th International Conference of Pattern Recognition 2010

Publication date:
2010

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Bespalov, D., Dahl, A. L., & Shokoufandeh, A. (2010). Geometric Total Variation for Texture Deformation. In Proceedings of 20th International Conference of Pattern Recognition 2010 (pp. 4597-4600).

DTU Library
Technical Information Center of Denmark

Geometric Total Variation for Texture Deformation

Dmitriy Bespalov
Department of Computer Science
Drexel University
Philadelphia, USA
Email: dbespalo@cs.drexel.edu

Anders Dahl
DTU Informatics
Technical University of Denmark
Lyngby, Denmark
Email: abd@imm.dtu.dk

Ali Shokoufandeh
Department of Computer Science
Drexel University
Philadelphia, USA
Email: ashokouf@cs.drexel.edu

Abstract—**In this work we propose a novel variational method that we intend to use for estimating non-rigid texture deformation. The method is able to capture variation in gray-scale images with respect to the geometry of its features. Accurate localization of features in the presence of unknown deformations is a crucial property for texture characterization. Our experimental evaluations demonstrate that accounting for geometry of features in texture images leads to significant improvements in localization of these features, when textures undergo geometrical transformations. In addition, feature descriptors using geometrical total variation energies discriminate between various regular textures with accuracy comparable to SIFT descriptors, while reduced dimensionality of TVG descriptor yields significant improvements over SIFT in terms of retrieval time.**

I. Introduction

In most natural images texture is an important element and contains information about the depicted scenes, for example information about object shape or class. As a consequence, texture analysis is a widely studied topic in image processing. We are interested in investigating the problem of characterizing textures that undergo non-rigid deformations. The primary goal of our research is to tackle the problem of estimating unknown non-rigid deformations in textures that exhibit some regularity, see Figure 1 for samples of such textures.

Intuitively, the spatial distribution of image features will determine how a texture is deformed, but the same features are characteristic for a texture. This makes characterization and deformation estimation dependent on each other. By image features we mean discontinuities in intensity values of the image plane, see e.g. [1]. Furthermore, if a gray-scale image is treated as $2D$ manifold (two spatial dimensions) embedded into $3D$ space (intensity becomes third dimension), then any single point placed on a surface of the image manifold will have co-dimensionality zero. In addition, two types of discontinuities (i.e. features) may be present in image manifold: features of co-dimension one and two. It is easy to see that discontinuities of co-dimension one will correspond to edges, while corners will be features of co-dimension two.

In general, due to dependency, simultaneously estimating the deformation and characterizing texture is difficult. While it is known that Total Variation (TV) energy is invariant under minor deformations [1], accurately estimating the geometry of deformed texture features via total variation energy is computational costly and will require some sort of flow estimation. In contrast, Total Variation Geometry (TVG) directly accounts for the intrinsic geometry of image features in the form of co-dimensionality, which significantly improves accuracy of tasks such as feature localization in the presence of deformation - a crucial property for both texture characterization and deformation estimation. In addition, we also show that deforming area of integration while computing the TVG based on the local geometry of texture, further improves localization of features. When computing TVG energy responses, we will refer to the process of deforming the region of integration as "steering TVG".

II. Related Work

A survey of texture characterization is given in for example [2], [3], where the latter has particular focus on deformation-invariant texture characterization. Often used categories for texture analysis are statistical, model based, and structural methods, which we will briefly describe here.

Statistical methods are concerned with the relative distribution of pixel intensities, and a distinction is made between first, second and higher order statistics [4], [2], [3]. First order is basically a histogram of intensities. Second order is the probability of co-occurrence of intensities, and early work of Haralick *et al.* [5] introduced the concept of gray level co-occurrence matrices. Higher order statistics concerns the joint probability of more than two pixels.

Model based methods are concerned with estimating the coefficients of a texture model and using the coefficients for texture characterization. Texture models include Markov random fields, Gabor filters, wavelet models, steerable filters, etc. Typically, a modeled texture can be generated by a linear combination of weighted basis functions. Different approaches for obtaining deformation invariance have been proposed, for example obtaining rotation or scale invariance by combining coefficients, which only differ by rotation or scale [3]. Recent examples of model based texture analysis include steerable filters and Markov random fields used in
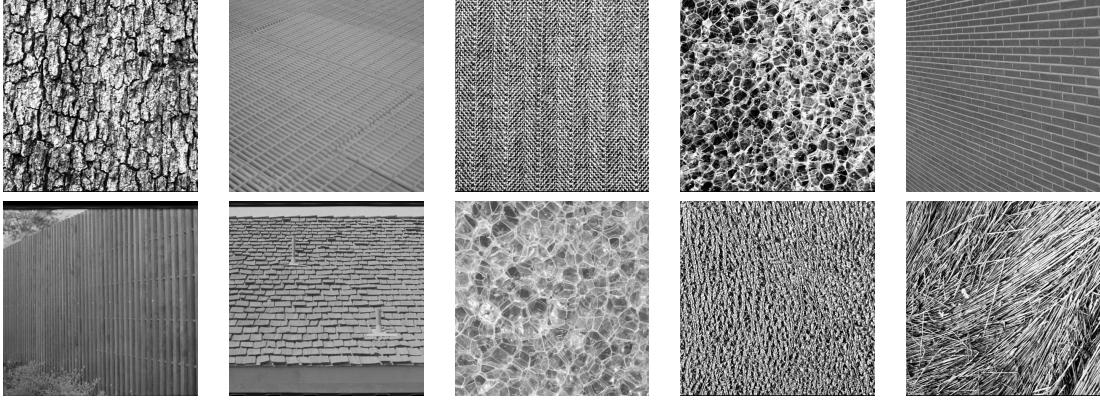
Figure 1. Textures used in experimental results.

[6] for estimating texture deformation, and TV energy used in [7] for separating textured and structure regions.

Structural methods are based on basic image structures like textons [8] or features obtained from image derivatives. Deformation is estimated in [9] and [10] employing a high level vision model, but based on texton-like structures. A low level computer vision approach is used in [11] based on image derivatives for estimating texture deformation in scale space. [12] suggest a sparse texture characterization based on size and orientation of image regions and local features. The spatial extension of the image regions are found from TV-flow. No existing feature based texture characterization explicitly account for the intrinsic geometry of the texture. We suggest a method based on Total Variation Geometry (TVG) presented in [1] that preserves feature co-dimensionality. It is worth mentioning that differential manifolds utilized in the estimation of the proposed Beltrami operator in [13] is closely related to TVG energy density. We will now explain the details of our method.

III. GEOMETRICAL TEXTURE CHARACTERIZATION

Geometrical Total Variation energy density (TVG) was proposed by Burchard in [1], and is defined as follows. Given a map u between manifolds $M_1 \rightarrow M_2$, TVG energy of degree r is the elementary symmetric function of degree r in the singular values of the derivative ∇u. Any gray-scale image, may be treated as a map $u : \Re^2 \rightarrow \Re^3$, in other words, image is a $2D$ manifold embedded into $3D$ space. As a result, for gray-scale images only two degrees of TVG energy density is available. First degree of TVG responds to discontinuities of co-dimensions one and two; second degree of TVG smoothes discontinuities of co-dimension one (edges) and responds to discontinuities of co-dimension two (corners).

We now briefly describe the method to compute TVG energy density for gray-scale images, which directly follows from theoretic results presented in [1]. Given image $I(x, y)$, let I_x and I_y denote first order derivatives of I along x and

y directions, respectively. And, TVG energy of I of orders one and two are $TVG_1(x, y) = \sigma_1 + \sigma_2$ and $TVG_2(x, y) = \sqrt{\sigma_1 \times \sigma_2}$, where σ_1 and σ_2 are singular values of matrix $\nabla I(x, y)$ filled with I_x and I_y for pixels in the neighborhood Ω of the pixel (x, y) (size of matrix $\nabla I(x, y)$ is $|\Omega| \times 2$).

Accurate estimation of first order derivatives of images is essential to stability of TVG energy density, especially if images are deformed with some geometrical transformations. We employ second order kernel regression to compute values of I_x and I_y described in [14], that formulates second order kernel regression as a weighted least-squares optimization problem. The weights are computed using Gaussian kernels with steering covariance matrices, so each neighbor $p_i = (x_i, y_i)^T$ of pixel $p = (x, y)^T$ is weighted by $K_{H_i}(p_i - p) = \exp(-(p_i - p)^T C_i(p_i - p)/2h)$, where C_i is a covariance matrix defined in [14], and h is a smoothing parameter.

The procedure to estimate image gradients is as follows. First, I_x and I_y are estimated using second order steering kernel regression. Second, steering covariance matrices C_i are computed using new estimates of image gradients. Steering covariance matrices C_i are initialized with identity matrices. We found that two iterations of the procedure is sufficient to produce robust estimates of image gradients, even when significant deformation is introduced to images.

The response of geometrical total variation energy greatly depends on the choice of neighbor set Ω. We propose to use steering Gaussian weights $K_{H_i}(p_i - p)$ to construct neighbor sets Ω_i for every pixel p_i. In other words, we would like to adapt shapes of integration regions Ω_i when computing TVG energies, based on local geometry of the texture. We refer to the process of estimating TVG energies in this way as steering TVG. For the sake of implementation efficiency we simulate the process of building Ω_i by multiplying image gradients in $\nabla I(p)$ with weights $K_{H_i}(p_i - p)$, while rows $(I_x(p_1), I_y(p_1))$ in $\nabla I(p)$ are selected from fixed size set Ω which corresponds to square window centered around pixel p. Illustrations of steering TVG for selected texture images

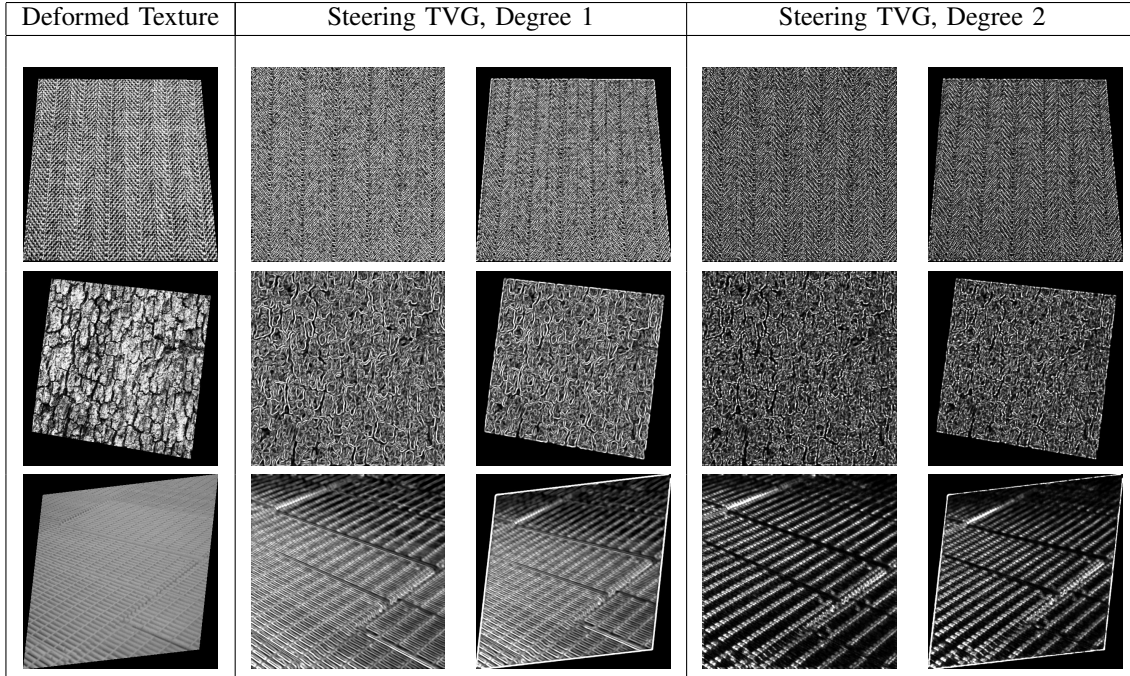| Deformed Texture	Steering TVG, Degree 1	Steering TVG, Degree 2

Figure 2. Geometrical total variation energy for selected images. Illustrations of steering TVG values are presented for original and deformed textures.

under various perspective transformations can be found in Figure 2.

A high dimensional feature descriptor can then be computed using the TVG energies. Like SIFT features [15], the TVG descriptor is defined by the position in the image, the scale and orientation of the feature, and a vector describing the feature. The image interest points are found in the same manner as SIFT features, i.e. position, scale, and orientation are similar to the SIFT feature. To create the feature vector, the TVG energies $TVG_1(x, y)$ and $TVG_2(x, y)$ are computed for the entire image. The scale, σ, of the feature is defined as standard deviation for the Gaussian kernel that corresponds to the local extrema of the scale-normalized Laplacian of Gaussian. Similar to [16], we set the size of the window for the feature descriptor to $20 \times \sigma$. We then rotate this square window centered around the interest point to match the computed feature orientation. The window is then divided evenly into 4×4 sub-windows in order to preserve spatial information.

The first and second degree TVG energies are summed for each sub-window to form part of the feature vector. To approximate the summation values at each sub-window we integrate TVG energies within axis-aligned windows of size $\frac{1}{2} \times \sigma$ using integral images, regularly sampled at 10×10 points. The feature descriptor is then defined by the concatenation of the TVG energies of each sub-window, resulting in a $4 \times 4 \times 2 = 32$ dimensional feature descriptor. Finally, the resulting feature vector is normalized to unit length to preserve invariance to changes in light intensity.

IV. EXPERIMENTAL RESULTS

In this work we investigate the accuracy of localization of features for three different types of variational energies: 1.) geometrical total variation with steering (steering TVG), 2.) geometrical total variation without steering (TVG), and 3.) total variation (TV). Recall that TV energy for images is defined as sum of gradients' ℓ_2 norms in a neighborhood Ω. In addition, we demonstrate that steering TVG feature descriptors introduced in III can be used to discriminate between various regular textures deformed with perspective transformations. The accuracy of steering TVG descriptors is comparable to SIFT, while running time is on the order of magnitude lower for steering TVG descriptors over SIFT.

We have chosen 10 texture images shown in Figure 1 and applied three various perspective transformations to every image (see first column of Figure 2). Each image in the resulting dataset of 40 textures was 512×512 pixels. We then computed scale-space interest points for three types of energies, and every interest point from the non-deformed texture was warped to the deformed image coordinates, and pixel distance between warped interest point and closest interest point from the deformed texture was recorded (distances were computed with respect to points' locations). Table I presents mean and standard deviation of these distances for the three types of energies tested. Clearly, steering TVG of both degrees outperforms two other types of energies tested. This indicates that steering TVG produces more accurate localization for texture features than other types of variational energies.

	TV		TVG, Degree 1		TVG, Degree 2		Steering TVG, Degree 1		Steering TVG, Degree 2	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
	0.47	0.24	0.46	0.24	0.45	0.24	0.38	0.22	0.37	0.22

Table I
LOCALIZATION ERRORS FOR THREE TYPES OF VARIATION ENERGIES.

Descriptor	Mean	Std Dev	Time
SIFT	0.97	0.02	62m 24s
TVG	0.93	0.05	8m 29s

Table II
TEXTURE CLASSIFICATION.

The original 10 texture images from Figure 1 and 30 deformed textures were used in classification experiment. For each texture image, SIFT and steering TVG feature descriptors corresponding to scale-space interest points found in gray-scale images were formed. We then attempted to classify every interest point by finding its nearest neighbor in the database. For every texture image, the database contained all interest points from the remaining 39 images. We recorded an interest point as correctly classified if its nearest neighbor was from the same texture class. Table II lists mean and standard deviation of classification accuracies (number of correctly classified interest points divided by total number of interest points) for all images in the dataset. Running times for classification experiments are also provided. The experiments indicate that steering TVG descriptors improve retrieval time over SIFT at the expense of insignificant degradation in classification accuracy.

V. FUTURE WORK

In this paper we described a model for texture characterization that accounts for geometry of the textured features. We have identified several promising directions for our future work. First, we would like to improve localization of scale-space interest points through steering Gaussian weighting schema when building Laplatian pyramids. Second, we would like to extend definition of steering TVG energies to multi-band (color) images. Finally, we would like to revisit different TV regularization methods used for segmentation and classification, and see if the use of steering TVG improves their performance.

ACKNOWLEDGMENTS

REFERENCES

[1] P. Burchard, "Total variation geometry I: Concepts and motivation," *UCLA CAM Report*, vol. 2, no. 01, 2002.

[2] A. Materka and M. Strzelecki, "Texture analysis methods-a review," *Technical University of Lodz, Institute of Electronics, COST B*, vol. 11, pp. 9–11, 1998.

[3] J. Zhang and T. Tan, "Brief review of invariant texture analysis methods," *Pattern Recognition*, vol. 35, no. 3, pp. 735–747, 2002.

[4] M. Bharati, J. Liu, and J. MacGregor, "Image texture analysis: methods and comparisons," *Chemometrics and intelligent laboratory systems*, vol. 72, no. 1, pp. 57–71, 2004.

[5] R. Haralick, I. Dinstein, and K. Shanmugam, "Textural features for image classification," *IEEE Transactions on systems, man, and cybernetics*, vol. 3, no. 6, pp. 610–621, 1973.

[6] J. Chang and J. Fisher III, "Analysis of Orientation and Scale in Smoothly Varying Textures," in *Proceedings of the International Conference on Computer Vision*, 2009.

[7] J. Aujol, G. Gilboa, T. Chan, and S. Osher, "Structure-texture image decompositionmodeling, algorithms, and parameter selection," *International Journal of Computer Vision*, vol. 67, no. 1, pp. 111–136, 2006.

[8] B. Julesz, "Textons, the elements of texture perception, and their interactions," *Nature*, vol. 290, no. 5802, pp. 91–97, 1981.

[9] J. Hays, M. Leordeanu, A. Efros, and Y. Liu, "Discovering texture regularity as a higher-order correspondence problem," *Lecture Notes in Computer Science*, vol. 3952, p. 522, 2006.

[10] W. Lin and Y. Liu, "Tracking Dynamic Near-Regular Texture Under Occlusion and Rapid Movements," *Lecture Notes in Computer Science*, vol. 3952, p. 44, 2006.

[11] T. Lindeberg and J. Gårding, "Shape from texture from a multi-scale perspective," 1993.

[12] T. Brox and J. Weickert, "A TV flow based local scale measure for texture discrimination," *Lecture Notes in Computer Science*, pp. 578–590, 2004.

[13] R. Kimmel, N. A. Sochen, and R. Malladi, "On the geometry of texture," Proceedings of the 4th International conference on Mathematical Methods for Curves and Surfaces, Tech. Rep., 1996.

[14] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel Regression for Image Processing and Reconstruction," *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 349–366, 2007.

[15] D. Lowe, "Object recognition from local scale-invariant features," vol. 2, 1999, pp. 1150 –1157 vol.2.

[16] H. Bay, T. Tuytelaars, and L. V. Gool, "Surf: Speeded up robust features," in *In ECCV*, 2006, pp. 404–417.