AUTOMATIC IMAGE DECOMPOSITION

Kedar A. Patwardhan and Guillermo Sapiro

Department of Electrical and Computer Engineering and Digital Technology Center,
University of Minnesota, Minneapolis, MN-55455, USA.
{*kedar, guille*}@*ece.umn.edu*

ABSTRACT

The decomposition of an image into its primitive components, such as cartoon plus texture, is a fundamental problem in image processing. In [11, 16], the authors proposed a technique to achieve this decomposition into structure and texture. These two components are competing ones, and their proposed model has a critical parameter that controls this decomposition. In this paper we show how to automatically select this parameter, and demonstrate with examples the importance of this optimal selection.

Keywords: Image decomposition, structure, texture, bounded variation, parameter selection, inpainting.

1. INTRODUCTION

Natural images usually contain two main components, which are broadly classified as structure and texture. Structure is mainly the sharp edges in the image, which generally separate different objects, or are caused by differences in shading. Texture, on the other hand, is in general a repetitive pattern, or "oscillations." The decomposition of an image into two parts, where one image contains the structure and the other image contains the texture, is important in a number of different areas, such as copying just the texture in one image onto another image, segmentation of an image based upon the texture or DC gray-values, etc. This decomposition has been shown in [6] to be fundamental for image inpainting, the art of modifying an image in a non-detectable form [2, 3, 5, 7, 10].

Following work by Meyer [11], Vese and Osher proposed in [16] an algorithm to achieve this decomposition, this being the technique exploited in [6] for image inpainting (see also [1, 9, 12, 14] for other related decomposition approaches). As we will see bellow, there is a critical parameter in this decomposition that controls the competition between the structure and the texture components. This parameter was left to the user to select, in the original formulation. Our goal in this paper is to simplify the decomposition, and thereby the above mentioned applications, by designing a very simple procedure for the automatic selection of this critical parameter. The devised method has been motivated by the approach used in [13] (see also [4] for automatic local selections of critical parameters for this Total Variation, TV, model), and automatically finds the parameter for the best possible decomposition in terms of a given penalty equation. We should note that this is in sharp contrast with multiscale approaches such as those reported in [15], where the authors opt for selecting a set of parameters, thereby producing a set of decompositions instead of just an optimal one as suggested here.

We thank Marcelo Bertalmio, Vicent Caselles, and Liron Yatziv for conversations on image inpainting and Stanley Osher for feedback on image decomposition. This work was partially supported by a grant from the Office of Naval Research, the National Science Foundation, the Presidential Early Career Awards for Scientists and Engineers (PECASE), and National Science Foundation CAREER Award.

2. IMAGE DECOMPOSITION: BACKGROUND AND OUR APPROACH

In this section, we will first throw some light on the image decomposition approach proposed in [16] and [11]. The main ingredients of this decomposition are the total variation minimization of [13] for image denoising and restoration, and the space of oscillating functions introduced in [11] to model texture or noise. We should note that this decomposition is not just a low-pass/high-pass one, since both the structure and texture images contain high and low frequencies, but of different type.

Let $I : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a given observed image, $I \in L^2(\mathbb{R}^2)$. I could be just a noisy version of a true underlying image u, or could be a textured image, u then being a simple sketchy approximation or a cartoon image of I (with sharp edges). A simple relation between u and I can be expressed by a linear model, introducing another function v, such that $I(x, y) = u(x, y) + v(x, y)$. In [13], the problem of reconstructing u from I is posed as a minimization problem in the space of functions of bounded variation $BV(\mathbb{R}^2)$, [8], allowing for edges:

$$\inf_{u \in BV} \left\{ F(u) = \int |\nabla u| + \lambda \|v\|^2_{L^2},\ I = u + v \right\}. \quad (1)$$

where $\lambda > 0$ is a tuning parameter. The second term in the energy is a fidelity term, while the first term is a regularizing term, to remove noise or small details, while keeping important features and sharp edges. The selection of the optimal parameter λ for this model is dealt with in [4, 13].

In [11], Meyer proved that for small λ the model will remove the texture. To extract both the $u \in BV$ (a piecewise constant or cartoon representation of the image), and the v component as an oscillating function (texture or noise) from I, Meyer proposed the use of a different space of generalized functions, which is in some sense the dual of the BV space (and therefore, contains oscillations).

Inspired by Meyer, in [16], Vese and Osher devised and solved a variant of his model, making use of partial differential equations.

The following minimization problem is the one proposed in [16]:

$$\inf_{u,g_1,g_2} \Big\{ G_p(u,g_1,g_2) = \int |\nabla u| + \lambda \int |I - u - \partial_x g_1 - \partial_y g_2|^2 dxdy + \mu \Big[\int \Big(\sqrt{g_1^2+g_2^2}\Big)^p dxdy \Big]^{\frac{1}{p}} \Big\}, \quad (2)$$

where $\lambda, \mu > 0$ are tuning parameters, and $p \to \infty$. The first term ensures that the cartoon image $u \in BV(\boldsymbol{R}^2)$, the second term ensures that $I \approx u + \text{div}(g_1, g_2)$ ((g_1, g_2) are the two auxiliary functions used by Meyer to define the new norm for oscillations), while the third term is a penalty on the norm of $v = \text{div}(g_1, g_2)$. This follows the space for oscillations proposed in [11], where the texture image v is given by $\text{div}(g_1, g_2)$, with the pair (g_1, g_2) having minimal norm.

For $p = 1$, the corresponding Euler-Lagrange equations are [16]

$$u = I - \partial_x g_1 - \partial_y g_2 + \frac{1}{2\lambda}\text{div}\Big(\frac{\nabla u}{|\nabla u|}\Big), \quad (3)$$

$$\mu \frac{g_1}{\sqrt{g_1^2+g_2^2}} = 2\lambda \Big[\frac{\partial}{\partial x}(u - I) + \partial^2_{xx} g_1 + \partial^2_{xy} g_2\Big], \quad (4)$$

$$\mu \frac{g_2}{\sqrt{g_1^2+g_2^2}} = 2\lambda \Big[\frac{\partial}{\partial y}(u - I) + \partial^2_{xy} g_1 + \partial^2_{yy} g_2\Big]. \quad (5)$$

As can be seen from the examples in [16] , the minimization model (2) allows to extract from a given real textured image I the components u and v, such that u is a sketchy (cartoon) approximation of I, and $v = div(g_1, g_2)$ represents the texture or the noise (note once again that this is not just a low/high frequency decomposition). For some theoretical results and the detailed semi-implicit numerical implementation of the above Euler-Lagrange equations, see [16].

Our goal in this paper is to develop a simple and fast method to automate the selection of the tuning parameter λ so that we minimize the energy in (2). This parameter controls the competition between the structure part u and the texture part v, and if wrongly selected, it will leave texture in u and/or strong boundaries on v, thereby not achieving the main goal of the decomposition and making subsequent uses of it more difficult.

To achieve this goal, we use a method similar to the gradient projection method used in [13]. We first make an approximation resulting from the model proposed in [13], that is, at equilibrium state,

$$\sqrt{g_1^2(x,y) + g_2^2(x,y)} = \frac{1}{2\lambda}, \quad (6)$$

see [16].

At every iteration of implementing the above Euler-Lagrange equations (3,4,5), we assume that we have reached equilibrium. Hence, the second term contributing to the energy in (2) is close to zero. Let us call the new (minimizing) values of u, g_1 and g_2 by the names u_{new}, ${g_1}_{new}$ and ${g_2}_{new}$. Suppose that the new value of λ (say, λ_{new}) minimizes the energy. Then, in a method similar to the gradient projection method, we equate the energy calculated from the previous values with that calculated from the new values which are supposed to attain equilibrium. Thus,

$$\int |\nabla u_{new}| + \mu \int (\frac{1}{2\lambda_{new}})dxdy = \int |\nabla u| + \lambda \int |I - u - \partial_x g_1 - \partial_y g_2|^2 dxdy + \mu \int \Big(\sqrt{g_1^2+g_2^2}\Big)dxdy. \quad (7)$$

gives us a simple expression for computing λ_{new}, for the next iteration of the above Euler-Lagrange equations.

Although the assumptions mentioned above do not need to be made, and a similar computation can be done without assuming the approximation for $g_1^2 + g_2^2$ and without neglecting the second term in the variational formulation, making these assumptions simplifies the computation, and experimentally we haven't observed any significant improvement when the real values of these terms are considered. In addition to using the approximation (6), we tried two other approaches. Firstly, we simply substituted λ_{new}, u_{new}, ${g_1}_{new}$ and ${g_2}_{new}$ in equation (2) and equated this to the energy computed from the previous iteration, so as to get the value for λ_{new}. Secondly, we tried using the value of the residue at equilibrium from the model proposed by [16], in place of the second term in the equation (2). This also gives a simple expression for getting the next λ. Both of these methods did not produce desirable results, especially in comparison with (7).

Thus, the steps of the proposed algorithm are:

1. Solve equations 3, 4 and 5, obtaining u_{new}, ${g_1}_{new}$ and ${g_2}_{new}$.
2. Compute λ_{new} from equation 7.
3. Assign new values to the old values, i.e., $u_{new} \to u$, ${g_1}_{new} \to g_1$ and ${g_2}_{new} \to g_2$.
4. Repeat steps 1, 2 and 3 for about 100 iterations [16].

3. EXPERIMENTAL RESULTS

The results shown next have been generated using the 'Image Processing Toolbox' in MATLAB©. Since there is just an addition of one simple step for getting the new λ at every iteration, there is very little overhead and the decomposition takes the same amount of time as in the case of [16]. Figure 1 shows the decomposition of the original image (top), and the corresponding cartoon (u) and texture (v) images, when an arbitrary value of λ was used for the decomposition. Notice that there is still a clear component of the brick texture left in the cartoon image which was not transferred completely to the texture image. The use of equation (7) leads to a much smoother and piecewise constant cartoon image as described in Figure 2. The optimal lambda used also gives a much smaller value for the energy in Equation (2), as shown in the plot on the left. In Figure 3 we show that arbitrary selections of λ can affect not only the u image as in Figure 1, but the v image as well. The slight light variation in the original image is passed-on to the v (texture) image with arbitrary λ, while this is eliminated with our algorithm. Lastly, in Fig.4, we describe the importance of automatic selection of the best λ by an illustrative example from the field of inpainting. The left figure was decomposed and inpainted as in [6], by using an arbitrary λ (center) and with the optimal λ

(right). The region pointed to by the arrow shows the discrepancies in inpainting caused due to a non-optimal decomposition.

4. CONCLUDING REMARKS

In this note we have shown how to automatically compute the critical parameter in the image decomposition model proposed in [11, 16]. The technique is very simple and adds virtually no computational cost to the overall decomposition procedure. We should note that we can also consider finding a piece-wise space varying λ, following work done in [4] for the TV model in [13]. Preliminary experiments in this direction show that although there is a clear improvement in the energy, visually there are no significant changes when compared with the global optima computed with the technique proposed here. Additional results in this direction will be reported elsewhere.

5. REFERENCES

[1] J. F. Aujol, G. Aubert, L. Blanc-Feraud, and A. Chambolle, "Image decomposition: Application to textured images and SAR images," *INRIA Sophia Antipolis Report* **RR-4704**, January 2003.

[2] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera, "Filling-in by joint interpolation of vector fields and grey levels," *IEEE Trans. Image Processing* **10**, pp. 1200-1211, August 2001.

[3] M. Bertalmio, A. L. Bertozzi, and G. Sapiro, "Navier-Stokes, fluid dynamics, and image and video inpainting," *Proc. IEEE Computer Vision and Pattern Recognition (CVPR)*, Hawaii, December 2001.

[4] M. Bertalmio, V. Caselles, B. Rouge, and A. Sole, "TV based image restoration with local constraints," *Journal of Scientific Computing* **19**, 2003.

[5] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," *Computer Graphics (SIGGRAPH 2000)*, pp. 417–424, July 2000.

[6] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher, "Simultaneous structure and texture image inpainting," *IEEE Trans. Image Processing* **12**, pp. 882-889, 2003.

[7] T. Chan and J. Shen, "Local inpainting models and TV inpainting," *SIAM J. Appl. Math.* **62:3**, pp. 1019-1043, 2001.

[8] L.C. Evans and R.F. Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press, London, 1992.

[9] K. Karu, A. K. Jain, and R. M. Bolle, "Is there any texture in the image?," *Pattern Recognition* **29:9**, pp. 1437–1446, 1996.

[10] S. Masnou and J. Morel, "Level-lines based disocclusion," *IEEE Int. Conf. Image Processing*, October 1998.

[11] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, AMS University Lecture Series **22**, 2002.

[12] S. Osher, A. Sole, and L. Vese, "Image decomposition and restoration using total variation minimization and the H^{-1} norm," *SIAM Multiscale Modeling and Simulation* **1**, pp. 349 - 370, 2003.

[13] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, no. 60, pp. 259–268, 1992.

[14] J.-L. Starck, M. Elad, and D. L. Donoho, "Image decomposition: Separation of texture from piece-wise smooth content," *Proc. SPIE Annual Meeting*, San Diego, California, August 2003.

[15] E. Tadmor, S. Nezzar, and L. Vese, "A multiscale image representation using hierarchical (BV, L^2) decompositions," *UCLA CAM Report* **03-32**, July 2003

[16] L. Vese and S. Osher, "Modeling textures with total variation minimization and oscillating patterns in image processing," *Journal of Scientific Computing* **19**, pp. 553-572, 2003.

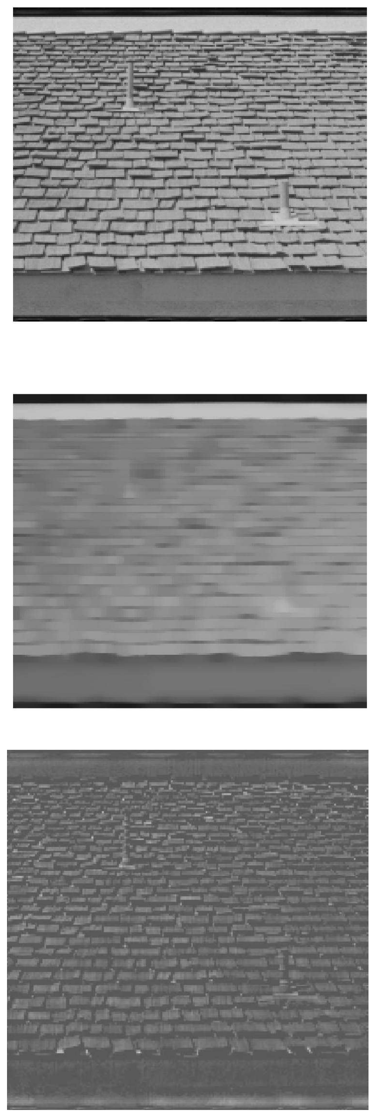

Fig. 1. Original image (top) and the corresponding decomposition. The cartoon image (center) has some texture still intact.

Fig. 2. Plot (left) shows how our method fares better than an arbitrary λ. The cartoon image (center) is more smooth and piecewise constant than the corresponding one in Figure 1.

Fig. 3. The image on the left is decomposed with arbitrary λ to produce the v image on the center and with optimal λ to produce the one on the right. The v image in the center still contains edges at the place of light contrast change (and not just oscillations), and these are gone with the λ computed with our approach. .

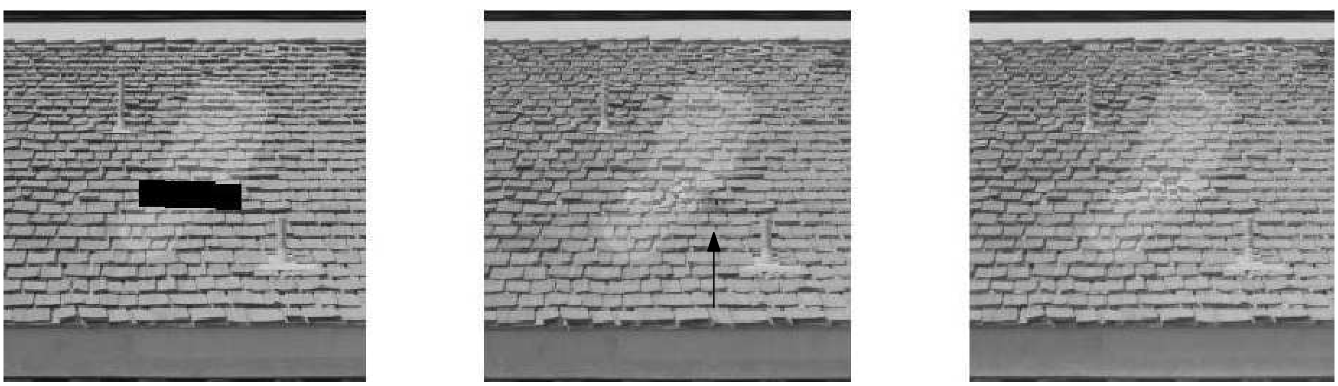

Fig. 4. Illustration of the application of our method for inpainting, following [6]. The image in the center is reconstructed using arbitrary λ and the one on the right uses our method for selecting λ.