Pacific Graphics (2015)
N. J. Mitra, J. Stam, and K. Xu (Editors) *Short Paper*

Structure-Preserving Image Smoothing via Phase Congruency-aware Weighted Least Square

Jinze Yu[1] and Yoichi Sato[1]

[1]Institute of Industrial Science, the University of Tokyo, Japan

Abstract

Structure-preserving image smoothing, or also understood as structure-texture separation problem, is an important topic for both computer vision and computer graphics as structure-texture separation can help better image understanding. In fact, many image processing problems can be well achieved once two layers possessing different properties of a scene are separated. Therefore better separating structure and texture from an image is of great practical importance. However, it is also a challenge topic since it is often quite subjective to tell the difference between the two layers. Recently, researchers made great efforts on separating a given image into its structure and texture layers by distinguishing edges from oscillations based on non-gradients-based descriptors or descriptors defined specifically for certain kinds of image data. These methods show advantages compared to the purely gradients-based methods with extra information provided besides gradients. In this paper, we propose a structure-texture separation method using non-gradients-based descriptor. Specially, we propose an alternative yet simple image smoothing approach based on the well-known weighted least square (WLS) framework. Our approach combines the phase congruency features that can better help locate structure or contour information of objects. Phase congruency performs well for distinguishing the structure and texture as it mimics the response of the human perception system to contours and is also sensitive to periodic patterns. By including the phase congruency as weights, WLS can better smooth out images while preserving structures. Experimental results indicate that the proposed approach is effective for structure-texture separation and achieves low computational complexity, compared to the state-of-the-art methods.

Categories and Subject Descriptors (according to ACM CCS): I.4.3 [Image Processing and Computer Vision]: Enhancement—Smoothing

1. Introduction

Natural scenes and human-created art pieces typically contain rich texture as can be seen on the handkerchief with a cartoon figure shown in Figure 1. Some natural scenes contain even more complicated and various patterns of texture (see Figure 3). While the human perception system can easily distinguish structure and texture inside images, understanding and then performing this structure-texture separation task poses great challenges for a computer. For example, it is difficult to discriminate fine-scale edges and details, both of them appear as small variation in the 1D signal domain even though the details have often appear as quasi-periodic. Previously, some researchers made great efforts into tackling this challenging problem of structure-texture separation [Yve01,SSD09,FFL10,BLMV10,XYXJ12,KEE13].

Figure 1: *An example of our structure-texture separation result.*

In fact, structure-texture decomposition can be formulat-

DOI: 10.2312/pg.20151274

ed as an estimation problem in which a given image is separated into two components that correspond to coarse and fine scale image details, respectively. Gaussian filter is the earliest and the most commonly used isotropic smoothing operator [Wit84, BA83]. Edge-aware smoothing approaches such as the use of Anisotropic diffusion filter [PM90], Total Variation model [ROF92], Bilateral filter [TM98, DD02], NL-means filter [BCM05], WLS filter [FFLS08] and L_0 smoothing [XLXJ11] utilize differences in intensity or color values or gradient magnitudes for predicting existence of edges, and then use the edge information to guide the smoothing process.

Such intensity variation or gradient-based definition of edges, however, might fail to capture high-frequency or periodic patterns that are related to fine image details or textures. Therefore these approaches cannot fully separate textured regions from the main structures as the edge indicators will consider such texture as part of the structure to be retained due to their large gradient magnitudes.

Subr *et al*. [SSD09] framed the decomposition problem in terms of local extrema modulation based on the fact that edges are determined by intensity oscillations between local extrema. Later on, Xu *et al*. [XYXJ12] proposed a relative total variation descriptor to better classify structure and texture elements, and they then proposed to include this information into the total variation framework to obtain better separation results. Karacan *et al*. [KEE13] adopted the region covariances to the non-local means filter and used it for image smoothing. By using region covariances commonly used for representing textures, their method is able to remove small scale textures from images while preserving structures. Zang *et al*. [ZHZ14] uses local extrema for feature characterization as Subr *et al*. [SSD09] did, but introduced the curvalization techniques to represent the 2D regions' property into 1D curve in order to reduce the 2D computation into 1D processing to achieve faster processing.

In this paper, we present a novel approach for structure-preserving image smoothing, also understood as structure-texture separation which we mention above, based on the weighted least squares (WLS) framework [FFLS08]. WLS has been shown to posses the nice property of smoothing image details at different scales without blurring the edges. However, we show that the original gradient-based WLS is not suitable for texture removal, while our proposed method based on phase congruency turns out to be a better alternative. This is because phase congruency can better represent the human visual system's response to contour and can detect periodic pattern since it works in the frequency domain [MRBO86, MO87, PP11]. Based on this observation, we employ a local phase-based measure to extract the structure map from images. The edge map is then incorporated into the WLS framework as weighting function to guide the optimization during smoothing. Experimental results show that our method achieves good performance and has low computational complexity, compared to the state-of-the-art methods. As demonstrated in Figures 1 and 3, the proposed model can effectively eliminate texture without distorting structure.

2. Our Method

2.1. Edge Detection Measure

It has been shown in the previous study that phase information gives evidence of object's contour [PP11]. In fact, the local energy model developed in [MRBO86] [MO87] postulates that features are perceived at points where the Fourier components are maximally in phase. According to the local energy model [MRBO86], points of maximum phase congruency (*e.g.*, 0° phase congruency at a step edge, or 90° phase congruency at a delta edge) could be the points where the visual system perceives a feature. Accordingly, based on the monogenic signal [FS01] (which is defined for the whole image), a local monogenic phase-based measure is defined for a given image to detect asymmetric features. The monogenic signal is defined by combining the 2D signal f with its Riesz transform (details can be found in [FS01]) $\mathbf{f}_R$ to form

$$\begin{aligned} f_M(x,y) &= (f,\mathbf{f}_R)(x,y) \\ &= (f(x,y),\mathbf{f}_R(x,y)) \\ &= (f(x,y),h_1 * f(x,y),h_2 * f(x,y)), \end{aligned} \quad (1)$$

where h_1 and h_2 are the Riesz filters, and x and y are 2D image coordinates, $*$ is the convolution operator. The definition of h_1 and h_2 is as follows:

$$h_1(x,y) = \frac{-x}{2\pi(x^2+y^2)^{\frac{3}{2}}},\ h_2(x,y) = \frac{-y}{2\pi(x^2+y^2)^{\frac{3}{2}}} \quad (2)$$

In practical applications, the local properties are analysed via several pairs of bandpass quadrature filters tuned to various spatial frequencies because real images generally consist of a wide range of frequencies. Therefore, a set of bandpass filters $c(x,y;s)$ (s is the scale parameter) are combined with the monogenic signal, which becomes

$$f_{M,s} = (c*f, c*h_1*f, c*h_2*f) = (even, \mathbf{odd}), \quad (3)$$

where functions *even* is $c*f$ and **odd** is $(c*h_1*f, c*h_2*f)$ They represent the scalar-valued even and the vector-valued **odd** responses of the quadrature filters, respectively. In stead of log-Gabor kernels used in [MRBO86], Cauchy kernels are adopted as a bandpass filter due to their good behaviour of localization. In the frequency domain, a 2D isotropic Cauchy kernel is defined as

$$C(\omega) = |\omega|^a \exp(-s|\omega|), \quad (4)$$

where $a \geq 1$, $\omega = (u,v)$, s is the scale parameter (same as in Eq. 2). Note that here C is the Fourier transform of function c (whose equation in spatially domain can be found in [BNB04]) in Eq. 3. More details about the parameters can be referred to [BNB04]. In all of our experiments, s is taken to be 13, and a is taken to be 1.5.

According to [Kov97], the absolute values of odd symmetric filter responses are large while the absolute values of even symmetric filter responses are small at points of asymmetry. The following measure is defined by using the differences between the odd and the even symmetric filter responses to detect asymmetric features.

$$FA = \sum_s \frac{\lfloor |\mathbf{odd}|_s - |even|_s - T_s \rfloor}{\sqrt{\mathbf{odd}_s^2 + even_s^2 + \varepsilon}}, \tag{5}$$

where ε is a small constant to avoid division by zero, T_s is the scale specific noise threshold, $|\cdot|$ is the ℓ^1-norm for **odd** and absolute value for *even*, $\lfloor\cdot\rfloor$ denotes the zeroing of negative values. The FA takes values in [0, 1], and is close to 0 in smooth regions and close to 1 near boundaries.

Some examples of FA edge map can be seen in Figure 2. It can be observed that edge maps generated using phase congruency correspond well to the manually created ground truth of structure map provided by Xu *et al*. [XYXJ12]. Unlike the gradient edge maps which contain a lot of edges due to texture, the edge maps from phase congruency capture the object contours while effectively suppressing edges from periodic patterns.

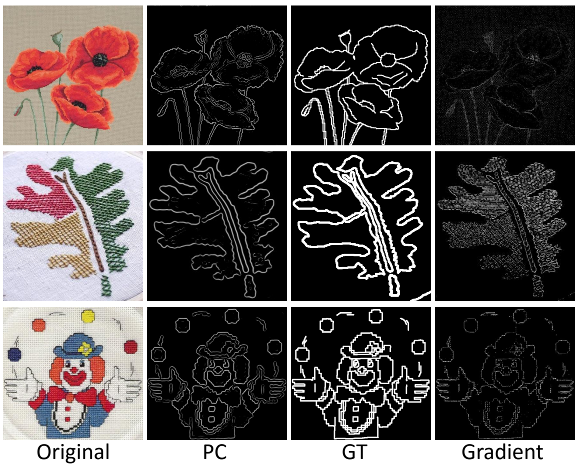

Figure 2: *Comparisons between phase congruency (PC), ground truth (GT) and gradient map.*

2.2. Phase Congruency Weighted Least Squares

To extract a structure layer from a natural image I, we aim to find a new image S which is as close to I as possible but is also as smooth as possible everywhere, except when passing across significant features. The weighted least squares (WLS) framework has been shown to perform well in smoothing image details while preserving edge features [FFLS08]. WLS carries out edge-aware smoothing by minimizing the following energy function

$$\sum_p \left((I_p - S_p)^2 + \lambda \left(w_p(I) \left(\frac{\partial S}{\partial x} \right)_p^2 + w_p(I) \left(\frac{\partial S}{\partial y} \right)_p^2 \right) \right) \tag{6}$$

where p represents image pixels and λ influences the smoothness of the optimized result S; w_p is a weight at the pixel p for the given image I.

However, WLS does not work well when directly applied to the structure-texture separation task because of the influences of texture. One major reason is that the weighting function are usually defined based on image gradients, and rather weak for indicating structure or contour but easily influenced by texture details, as shown in Figure 2. We solve this problem by combining the phase congruency-based FA (Eq. 5) into the WLS framework. Specifically, we develop a phase congruency weighted least squares (PCWLS) framework by setting the weighting function w_p as

$$w_p(I) = ((FA_p(I))^{\alpha} + \varepsilon')^{-1} \tag{7}$$

where α controls the sensitivity of the FA edge map (see Eq. 5) and its value can be set to be between 1 and 2; ε' is added to avoid dividing by zero. In our experiment, α is set to 1.5. As is demonstrated in our experiments, w_p in Eq. 7 take large effect in adapting WLS to structure-preserving image smoothing. Different from gradient-based operators, the FA measure is sensitive to structure edges and less sensitive texture edges. The smoothing process of WLS is prohibited near to the structure but greatly encouraged in homogeneous regions or regions with rich textures, resulting in smoothed images with structures being preserved.

Eq. 6 can be rewritten into the following matrix form

$$(I-S)^T(I-S) + \lambda(S^T D_x^T W D_x S + S^T D_y^T W D_y S) \tag{8}$$

where D_x and D_y are the Toeplitz matrices formed by arranging the discrete gradient operators with forward difference according to [XYXJ12]. W is diagonal matrix whose values are set as $W(i,i) = w_{p_i}(I)$, where i is the order of pixel p_i in the vector formed from I. Finally, the optimization of Eq. 6 can be written into the following analytical sparse linear system

$$(J + \lambda D_x^T W D_x + D_y^T W D_y)S = I \tag{9}$$

where J is the identity matrix. Note that this system is almost the same as in [FFLS08] for solving the WLS optimization. Actually, in our method we use the same solver as there for solving our problem.

3. Experiments

In our experiments, we compared our approach with some state-of-the-art edge preserving smoothing methods: weighted least squares (WLS) [FFLS08], relative total variation (RTV) [XYXJ12], and region covariances-based nonlocal means filter (RCNLM) [KEE13] on two images from [XYXJ12] (the first and third rows in Figure 3) and one image from [KEE13] (the second row in Figure 3). Experiments were carried out on a PC equipped with an Intel i7-3610 QM 2.10 GHz CPU and 8GB memory. Our code was written in Matlab 2013b. The source codes of the compared methods were obtained from websites provided by the authors of [FFLS08,XYXJ12,KEE13]. For all the tested methods, their parameters were carefully chosen to achieve better results. All of the methods were qualitatively evaluated on the basis that a good method should only smooth fine details and textures and preserve structure. In addition, the extracted texture or so-called detail component should be devoid of any information regarding the structure. Comparison on computational complexity were also done and the results were listed in Table 1. In all of our experiments on the three images, the parameter λ was set to 0.01.

From Figure 3 and Table 1, we can see that our method achieved comparable or better separating of structure layers with less computational cost, compared to the other methods. To be specific, WLS yielded large color bleeding and blurring effect, even if it had the lowest runtime. Although RCNLM gave pleasing qualitative result, its runtime was significantly larger than the other methods. Finally, both RTV and our method could achieve good smoothed images, whereas our method ran a bit faster. Also, as seen in row 2 in Figure 3, some artifacts would be enhanced by RTV.

4. Conclusions

In this paper, we proposed to incorporate a phase congruency-based edge map as weights into the WLS optimization framework to address the problem of structure-preserving image smoothing. We have demonstrated that our method offers at-least comparable qualitative results compared with the state-of-the-art methods but with significantly reduced computational cost. A shortcoming of our method is that there may sometimes be missing structure edge information in the FA edge map thus result in artifacts and color bleeding. In addition, as it is the same for other methods, our method is not effective enough for low-contrast contour. Future research will be in the direction of adopting multi-scale strategies to solve the mentioned problems.

Table 1: *Computational Complexity of Images in Figure 3 of Different Methods. (numbers in brackets show the size)*

Method	row1 (1024×768)	row2 (710×511)	row3 (495×536)
WLS	7.0s	4.2s	2.9s
RTV	8.8s	4.9s	3.6s
RCNLM	1124.2s	553.4s	358.6s
Ours	7.6s	4.6s	3.4s

References

[BA83] BURT P., ADELSON E.: The laplacian pyramid as a compact image code. *Communications, IEEE Transactions on 31*, 4 (Apr 1983), 532–540. 2

[BCM05] BUADES A., COLL B., MOREL J.-M.: A non-local algorithm for image denoising. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* (June 2005), vol. 2, pp. 60–65 vol. 2. 2

[BLMV10] BUADES A., LE T., MOREL J.-M., VESE L.: Fast cartoon + texture image filters. *Image Processing, IEEE Transactions on 19*, 8 (Aug 2010), 1978–1986. 1

[BNB04] BOUKERROUI D., NOBLE J. A., BRADY M.: On the choice of band-pass quadrature filters. *Journal of Mathematical Imaging and Vision 21*, 1-2 (2004), 53–80. 2, 3

[DD02] DURAND F., DORSEY J.: Fast bilateral filtering for the display of high-dynamic-range images. *ACM Trans. Graph. 21*, 3 (July 2002), 257–266. 2

[FFL10] FARBMAN Z., FATTAL R., LISCHINSKI D.: Diffusion maps for edge-aware image editing. *ACM Trans. Graph. 29*, 6 (Dec. 2010), 145:1–145:10. 1

[FFLS08] FARBMAN Z., FATTAL R., LISCHINSKI D., SZELISKI R.: Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Trans. Graph. 27*, 3 (Aug. 2008), 67:1–67:10. 2, 3, 4, 5

[FS01] FELSBERG M., SOMMER G.: The monogenic signal. *IEEE Transactions on Signal Processing 49*, 12 (2001), 3136–3144. 2

[KEE13] KARACAN L., ERDEM E., ERDEM A.: Structure-preserving image smoothing via region covariances. *ACM Trans. Graph. 32*, 6 (Nov. 2013), 176:1–176:11. 1, 2, 3, 4, 5

[Kov97] KOVESI P.: Symmetry and asymmetry from local phase. In *Tenth Australian joint conference on artificial intelligence* (1997), vol. 190, Citeseer. 3

[MO87] MORRONE M. C., OWENS R. A.: Feature detection from local energy. *Pattern Recognition Letters 6*, 5 (Dec. 1987), 303–313. 2

[MRBO86] MORRONE M. C., ROSS J., BURR D. C., OWENS R.: Mach bands are phase dependent. *Nature 324*, 6094 (1986), 250–253. 2

[PM90] PERONA P., MALIK J.: Scale-space and edge detection using anisotropic diffusion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 12*, 7 (Jul 1990), 629–639. 2

[PP11] PAPARI G., PETKOV N.: Review article: Edge and line oriented contour detection: State of the art. *Image Vision Comput. 29*, 2-3 (Feb. 2011), 79–103. 2

[ROF92] RUDIN L. I., OSHER S., FATEMI E.: Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena 60*, 1ĺC4 (1992), 259 – 268. 2

[SSD09] SUBR K., SOLER C., DURAND F.: Edge-preserving multiscale image decomposition based on local extrema. *ACM Trans. Graph. 28*, 5 (Dec. 2009), 147:1–147:9. 1, 2

[TM98] TOMASI C., MANDUCHI R.: Bilateral filtering for gray and color images. In *Computer Vision, 1998. Sixth International Conference on* (Jan 1998), pp. 839–846. 2

[Wit84] WITKIN A. P.: Scale-space filtering: A new approach to multi-scale description. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '84.* (Mar 1984), vol. 9, pp. 150–153. 2

[XLXJ11] XU L., LU C., XU Y., JIA J.: Image smoothing via l0 gradient minimization. *ACM Trans. Graph. 30*, 6 (Dec. 2011), 174:1–174:12. 2

[XYXJ12] XU L., YAN Q., XIA Y., JIA J.: Structure extraction from texture via relative total variation. *ACM Trans. Graph. 31*, 6 (Nov. 2012), 139:1–139:10. 1, 2, 3, 4, 5

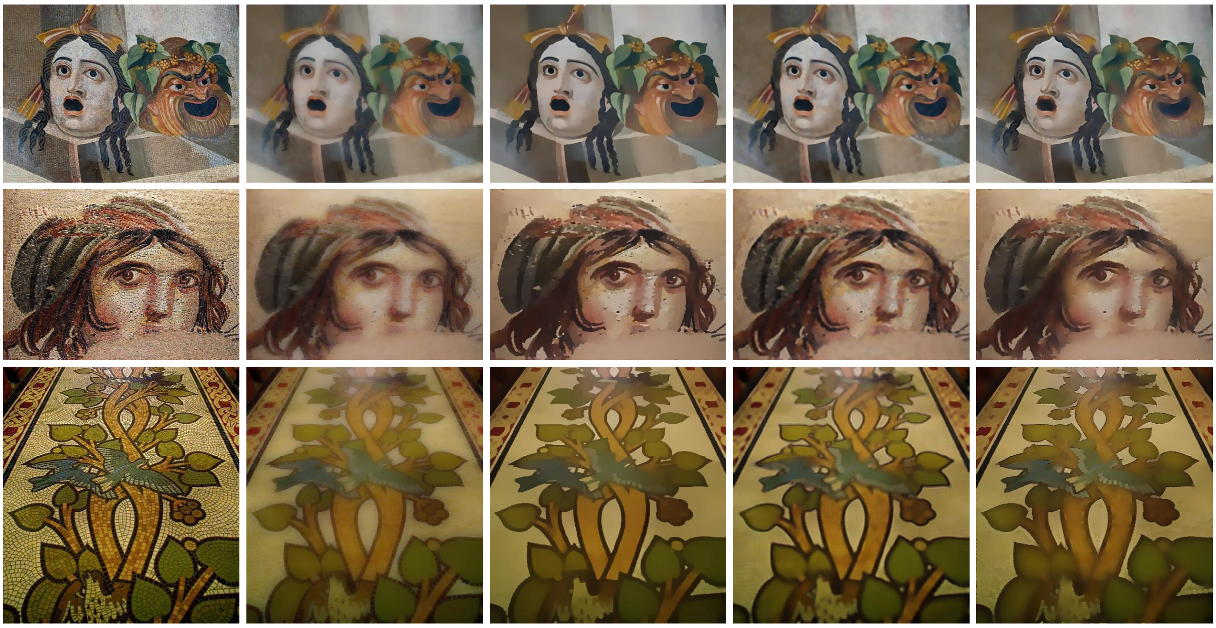

Figure 3: *Comparisons between our method (PCWLS), WLS [FFLS08], RTV [XYXJ12] and RCNLM [KEE13].*

[Yve01] YVES M.: Oscillating patterns in image processing and nonlinear evolution equations: The fifteenth dean jacqueline b. lewis memorial lectures. *Amer Mathematical Society 22* (2001). 1

[ZHZ14] ZANG Y., HUANG H., ZHANG L.: Efficient structure-aware image smoothingby local extrema on space-filling curve. *Visualization and Computer Graphics, IEEE Transactions on 20*, 9 (2014), 1253–1265. 2