

Dynamic Color Texture Modeling and Color Video Decomposition Using Bounded Variation and Oscillatory Functions

Mathieu Lugiez^{1,3}, Michel Ménard^{1,3}, and Abdallah El-Hamidi^{2,3}

¹ L3i - Université de La Rochelle

² MIA - Université de La Rochelle

³ Fédération PRIDES: Pôle régional de Recherche en Images, Données et Systèmes

Abstract. Dynamic, or temporal, texture is a spatially repetitive, time-varying visual pattern that forms an image sequence with a certain temporal stationarity. Important tasks are thus the detection, segmentation and perceptual characterization of Dynamic Texture (DT). Following recent work, color image decomposition appears as a good way to reach these different aims, however, to our best knowledge, no proposed model is currently able to deal with temporal aspect, inherent to color image sequences.

The major contribution of this paper is to adapt static decomposition model to time aspect in order to deal with videos and color image sequences. In this paper we propose an extended decomposition model which splits a video into two components, a first one containing geometrical information, the structure of the sequence and a second one dynamic color texture and noise. Examples for color video decomposition and characterization of real dynamic present in texture component will be presented.

1 Introduction

1.1 Motivation

A new issue in texture analysis is its extension to the temporal domain, a field known as Dynamic Texture (DT) analysis. In DT, the notion of self-similarity central to conventional image textures is extended to the spatiotemporal domain. DT are typically result from processes such as of water flows, smoke, fire, a flag blowing in the wind, a moving escalator, or a walking crowd. Important tasks are thus the detection, segmentation and perceptual characterization of dynamic textures. The ultimate goal is to be able to support video queries based on the recognition of the actual natural and artificial dynamic texture processes.

Following recent work, color image decomposition into geometrical and texture components appears as a good way to reach this aim in extracting meaningful information, i.e texture component, independently of geometrical information.

In this way, we propose to extend spatial color decomposition model to spatiotemporal domain, and attempt to highlight time influence present in video to characterize dynamic texture. To the best of our knowledge, no such time adaptation is currently available.

1.2 Overview of the Paper

The aim of this work is to extend a model, which decompose color image into two components, a first one containing geometrical structure U and a second one V , holding textural information and noise. So, we aim to deal with color image sequences in extending to time existing reliable model. Moreover, through decomposition of texture component, we attempt to determinate time impact, showing up dynamicity present in sequences, which will be suit for future work on dynamic texture.

In the first place of this paper we introduce the extended minimization functional problem and the associate discrete framework in which we place ourself and which is an appropriate one in image sequence processing. In a second part we present the extended time decomposition model and subsequently its color implementation and choice of parameters. Then, in the last part, we show some significant results and attempt to present time weight in dynamic texture extraction through color image sequence.

2 Time Extension of Decomposition Model

Decomposing an image into meaningful component appears as one of major aims in recent development in image processing. The first goal was image restoration and demising; but following the ideas of Yves Meyer, in total variation minimization framework of L. Rudin, S. Osher and E. Fatemi, image decomposition into geometrical and oscillatory (i.e texture) component appears as usefull and very interesting way in computer vision and image analysis. There is a very large literature and recent advances on image decomposition models, image regularization and texture extraction and modelling. So, we only cite, among many others, most recent works which we appear like most relevant and usefull paper. In this way, reader can refer to the work of Stark et al. [1], Aujol et al. [2], [3], Aujol and Chambolle [4], Aujol and Ha Kang [5] and Vese and Osher [6], [7], [8] to cover the most recent and relevant advances.

2.1 Decomposition Model and Functional Spaces

In order to decompose image sequences in suitable components we propose to extend the Osher-Vese [8] color decomposition model. Their approach derived from Meyer decomposition model [9]. Its rely on space of functions of bounded variation, BV and Banach space, G . Authors propose to minimize the following discretized functional:

$$\inf_{(u,v) \in BV \times G / f = u+v} \left\{ F(u, v) = \int |\nabla u| + \lambda \|v\|_G \right\} \quad (1)$$

BV is a good space to model functions with discontinuities along lines and curves, that can represent and preserve edges in an image. For the reader convenience we recall here its definition [10]:

Definition 1. $BV(\Omega)$ is the subspace functions $u \in L^1(\Omega)$ such that the following quantity, called the total variation of u , is finite:

$$J(u) = \sup \left\{ \int_{\Omega} u(x) \operatorname{div}(\xi(x)) dx \right\} \quad (2)$$

such that $\xi \in C_c^1(\Omega, \mathbb{R}^2)$, $\|\xi\|_{L^\infty(\Omega)} \leq 1$

In order to improve representation of oscillating patterns, Meyer in [9] has introduced the Banach space, G , which modeling signals with large oscillations, in particular textures and noise. A function belonging to G may have large oscillations and nevertheless have a small norm. Thus the norm on G is well-adapted to capture the oscillations of a function in an energy minimization method. We recall here the definition of G :

Definition 2. G is the Banach space composed of the distributions f which can be written $f = \partial_1 g_1 + \partial_2 g_2 = \operatorname{div}(g)$ with g_1 and g_2 in $L^\infty(\Omega)$. On G , the following norm is associate:

$$\|v\|_G = \inf \left\{ \|g\|_{L^\infty(\Omega, \mathbb{R}^2)} / v = \operatorname{div}(g), g = (g_1, g_2), |g(x)| = \sqrt{(|g_1|^2 + |g_2|^2)(x)} \right\} \quad (3)$$

2.2 Spatiotemporal Structure and Discretization

To take into account the spatiotemporal structure, we consider a video as an 3-D image [11], i.e a volume, so that we can apply 2-D image algorithms extended to the 3-D case. We assume that we have a given image sequence $f \in L^2(\Omega)$, where Ω is an open and bounded domain on \mathbb{R}^3 , with Lipschitz boundary. In order to recover u and v from f , we propose:

- An extended discrete version of gradient vector $|\nabla u|$ given by:

$$\begin{aligned} (\nabla u)_{i,j,k} &= ((\nabla u)_{i,j,k}^1, (\nabla u)_{i,j,k}^2, (\nabla u)_{i,j,k}^3) \\ (\nabla u)_{i,j,k}^1 &= \begin{cases} u_{i+1,j,k} - u_{i,j,k} & \text{if } i < N \\ 0 & \text{if } i = N \end{cases} \\ (\nabla u)_{i,j,k}^2 &= \begin{cases} u_{i,j,k} - u_{i,j+1,k} & \text{if } j < N \\ 0 & \text{if } j = N \end{cases} \\ (\nabla u)_{i,j,k}^3 &= \begin{cases} u_{i,j,k} - u_{i,j,k-1} & \text{if } k < N \\ 0 & \text{if } k = N \end{cases} \end{aligned} \quad (4)$$

- An extended discrete total variation definition:

$$J(u) = (\nabla u)_{i,j,k}^1 + (\nabla u)_{i,j,k}^2 + \alpha (\nabla u)_{i,j,k}^3 \quad (5)$$

we introduce the α constant to maintain homogeneity between space and time component. It's mainly for numerical implementation, to avoid discretization problem due to quantization step, which be different along space and time dimension. In practice, we often set it to one, but user can adapt it to less, more or in function of frame per second, or quickness of movement present in sequence, to ensure most reliability and homogeneity.

- An adapted definition of G , inspired by [8] for the vector case, extended to the third dimension:

Definition 3. Let G denote the Banach space consisting of all generalized vector-valued functions $\vec{v}(x, y, t) = (v_R(x, y, t), v_G(x, y, t), v_B(x, y, t))$ which can be written as:

$$\vec{v}(x, y, t) = (\operatorname{div} \vec{g}_R, \operatorname{div} \vec{g}_G, \operatorname{div} \vec{g}_B) \quad (6)$$

$$g_{1,c}, g_{2,c}, g_{3,c} \in L^\infty(\mathbb{R}^3), c = R, G, B,$$

induced by the norm $\|v\|_*$ defined as the lower bound of all L^∞ norms of functions $|\vec{g}|$, where $|\vec{g}| = \sqrt{|\vec{g}_R|^2 + |\vec{g}_G|^2 + |\vec{g}_B|^2} = \sqrt{\sum_{c=R,G,B} ((g_{1,c})^2 + (g_{2,c})^2 + (g_{3,c})^2)}$, and where the infimum is computed over all decompositions (6) of \vec{v} .

3 Description of the Extended Decomposition Model

For the reader convenience we only use, by now, useful indices in our formulas to avoid overloaded notation, so we will use c indices to denote color channel ($c \in R, G, B$) and 1, 2 and 3 exponents to represent the dimension to which objects belong.

3.1 Description of the Model

We propose the following minimization problem inspired by (6) and [8], for each color channel:

$$\inf_{u, g_1, g_2, g_3} \left\{ G_p(u, g_1, g_2, g_3) = \int |\nabla u| + \lambda \int |f - u - \partial_x g_1 - \partial_y g_2 - \partial_t g_3|^2 dx dy dt + \mu \left[\int (\sqrt{g_1^2 + g_2^2 + g_3^2})^p dx dy dt \right]^{\frac{1}{p}} \right\} \quad (7)$$

where λ, μ are tuning parameters (see next section for more useful details). Let reintroduce that $\vec{u} = (u_R, u_G, u_B)$, and $\vec{g}_i = (\vec{g}_{i,R}, \vec{g}_{i,G}, \vec{g}_{i,B})$, $i \in \{1, 2, 3\}$.

Formally minimizing the above energy equation with respect to u, g_1, g_2, g_3 , yields the following Euler-Lagrange equation for each color channel:

$$u = f - \partial_x g_1 - \partial_y g_2 - \partial_t g_3 + \frac{1}{2\lambda} \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) \quad (8)$$

$$\mu (\|\sqrt{g_1^2 + g_2^2 + g_3^2}\|)^{1-p} (\sqrt{g_1^2 + g_2^2 + g_3^2})^{p-2} g_1 = 2\lambda \left[\frac{\partial}{\partial x} (u - f) + \partial_{xx}^2 g_1 + \partial_{xy}^2 g_2 + \partial_{xt}^2 g_3 \right] \quad (9)$$

$$\mu (\|\sqrt{g_1^2 + g_2^2 + g_3^2}\|)^{1-p} (\sqrt{g_1^2 + g_2^2 + g_3^2})^{p-2} g_2 = 2\lambda \left[\frac{\partial}{\partial y} (u - f) + \partial_{yx}^2 g_1 + \partial_{yy}^2 g_2 + \partial_{yt}^2 g_3 \right] \quad (10)$$

$$\mu (\|\sqrt{g_1^2 + g_2^2 + g_3^2}\|)^{1-p} (\sqrt{g_1^2 + g_2^2 + g_3^2})^{p-2} g_3 = 2\lambda \left[\frac{\partial}{\partial t} (u - f) + \partial_{xt}^2 g_1 + \partial_{yt}^2 g_2 + \partial_{tt}^2 g_3 \right] \quad (11)$$

To simplify the presentation, let us introduce the notation:

$$H(g_1, g_2, g_3) = (\|\sqrt{g_1^2 + g_2^2 + g_3^2}\|)^{1-p} (\sqrt{g_1^2 + g_2^2 + g_3^2})^{p-2}$$

3.2 Numerical Implementation

For the reader convenience, we present here a numerical scheme to solve our minimization problem. So to solve equations (8) to (11) we use a semi-implicit finite differences scheme and an iterative algorithm, based on fixed point iteration [6].

We got, in order to compute gradient:

$$|\nabla u| \doteq |\nabla u|_{i,j,k} = \sqrt{\left(\frac{u_{i+1} - u_{i-1}}{2h}\right)^2 + \left(\frac{u_{j+1} - u_{j-1}}{2h}\right)^2 + \left(\frac{u_{k+1} - u_{k-1}}{2h}\right)^2} \quad (12)$$

so, belong first dimension (x) we got:

$$\begin{cases} |\nabla u|_{i+\frac{1}{2},j,k} = \sqrt{\left(\frac{u_{i+1} - u}{h}\right)^2 + \left(\frac{u_{j+1} - u_{j-1}}{2h}\right)^2 + \left(\frac{u_{k+1} - u_{k-1}}{2h}\right)^2} \\ |\nabla u|_{i-\frac{1}{2},j,k} = \sqrt{\left(\frac{u - u_{i-1}}{h}\right)^2 + \left(\frac{u_{i-1,j+1} - u_{i-1,j-1}}{2h}\right)^2 + \left(\frac{u_{i-1,k+1} - u_{i-1,k-1}}{2h}\right)^2} \end{cases} \quad (13)$$

in the same way, we compute gradient belong y and t dimension.

We are now in position to introduced complete numerical scheme to solving Euler-Lagrange equations:

$$u^{n+1} = \frac{1}{1 + \frac{1}{3\lambda h^2} \left(\frac{1}{|\nabla|_{i+\frac{1}{2}}} + \frac{1}{|\nabla|_{i-\frac{1}{2}}} + \frac{1}{|\nabla|_{j+\frac{1}{2}}} + \frac{1}{|\nabla|_{j-\frac{1}{2}}} + \frac{1}{|\nabla|_{k+\frac{1}{2}}} + \frac{1}{|\nabla|_{k-\frac{1}{2}}} \right)} * \left[f - \partial_x g_1 - \partial_y g_2 - \partial_t g_3 + \frac{1}{3\lambda h^2} \left(\frac{u_{i+1}^n}{|\nabla|_{i+\frac{1}{2}}} + \frac{u_{i-1}^n}{|\nabla|_{i-\frac{1}{2}}} + \frac{u_{j+1}^n}{|\nabla|_{j+\frac{1}{2}}} + \frac{u_{j-1}^n}{|\nabla|_{j-\frac{1}{2}}} + \frac{u_{k+1}^n}{|\nabla|_{k+\frac{1}{2}}} + \frac{u_{k-1}^n}{|\nabla|_{k-\frac{1}{2}}} \right) \right] \quad (14)$$

$$g_1^{n+1} = \frac{3\lambda}{\mu H(g_1^n, g_2^n, g_3^n) + \frac{6\lambda}{h^2}} \left[\partial_x u^n - \partial_x f + \frac{g_{1i+1}^n + g_{1i-1}^n}{h^2} + \partial_{xy}^2 g_2^n + \partial_{xt}^2 g_3^n \right] \quad (16)$$

$$g_2^{n+1} = \frac{3\lambda}{\mu H(g_1^n, g_2^n, g_3^n) + \frac{6\lambda}{h^2}} \left[\partial_y u^n - \partial_y f + \partial_{xy}^2 g_1^n + \frac{g_{2j+1}^n + g_{2j-1}^n}{h^2} + \partial_{yt}^2 g_3^n \right] \quad (17)$$

$$g_3^{n+1} = \frac{3\lambda}{\mu H(g_1^n, g_2^n, g_3^n) + \frac{6\lambda}{h^2}} \left[\partial_t u^n - \partial_t f + \partial_{xt}^2 g_1^n + \partial_{yt}^2 g_2^n + \frac{g_{3k+1}^n + g_{3k-1}^n}{h^2} \right] \quad (18)$$

$$(19)$$

where h denote the discrete step space, so in practice we set $h = 1$ in our algorithm, moreover in furthrance of time computing, we always use the most recent computed values. For the initialization of parameters we set: $u^0 = f$, $g_1^0 =$

$$\frac{-1}{3\lambda} \frac{\partial_x f}{\sqrt{\partial_x f^2 + \partial_y f^2 + \partial_t f^2}}, g_2^0 = \frac{-1}{3\lambda} \frac{\partial_y f}{\sqrt{\partial_x f^2 + \partial_y f^2 + \partial_t f^2}} \text{ and } g_3^0 = \frac{-1}{3\lambda} \frac{\partial_t f}{\sqrt{\partial_x f^2 + \partial_y f^2 + \partial_t f^2}}$$

3.3 Choice of Parameters

To obtain weak regularization, we use parameters as: $\lambda = 0.1$, $\mu = 0.1$ and we compute about 20 iterations. For strong regularization we got: $\lambda = 0.005$, $\mu = 0.005$ iterated 50 times. So, for classic decomposition we use parameters as: $\lambda = 0.01$, $\mu = 0.01$ iterated 20 times. For all results present in the next section we computed our decomposition on 8 images treated simultaneously as one block, but our algorithm is able to deal with bigger image block to catch lower frequencies and wider range of time oscillating patterns.

4 Numerical Results and Time Component

All images and results are compute from DynTex, the dynamic texture database [12] which provide a large and diverse database of high-quality dynamic textures. Dyntex sequences come from natural scene presenting a wide variety of moving process as flowing water, leaves blowing in wind, walking crowd... Such diversity grants user to identify and emphasize a lot of aspects in testing purpose.

For more details, demonstration sequence, wider range of results and for a presentation of similar color image sequences decomposition, relying on different approaches in the same framework [13], please consult this URL: <http://perso.univ-lr.fr/mlugiez>.

4.1 Static vs. Dynamic Decomposition

In order to prove that our dynamic decomposition method show more significant result than a static decomposition, we present a comparison between two

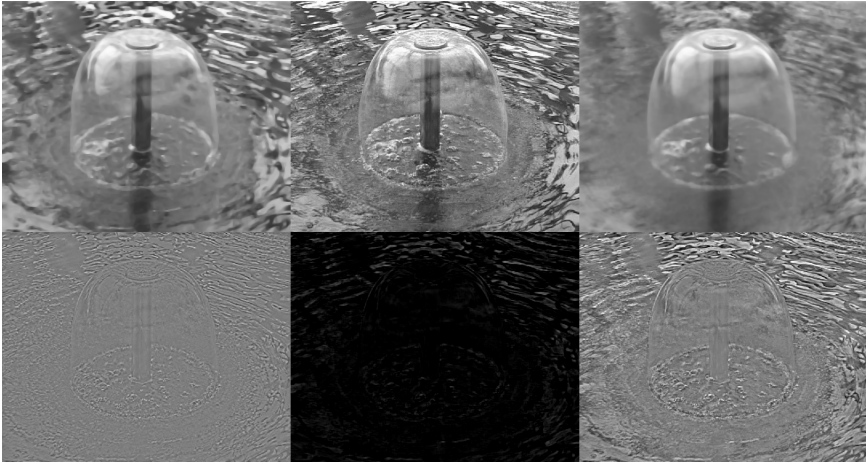


Fig. 1. From left to right, the geometrical component, U , in classic color decomposition (top) and its texture and noise component V (bottom). The original image (center and top), the difference between static and dynamic V component (center and bottom). Then our new dynamic decomposition components (right).

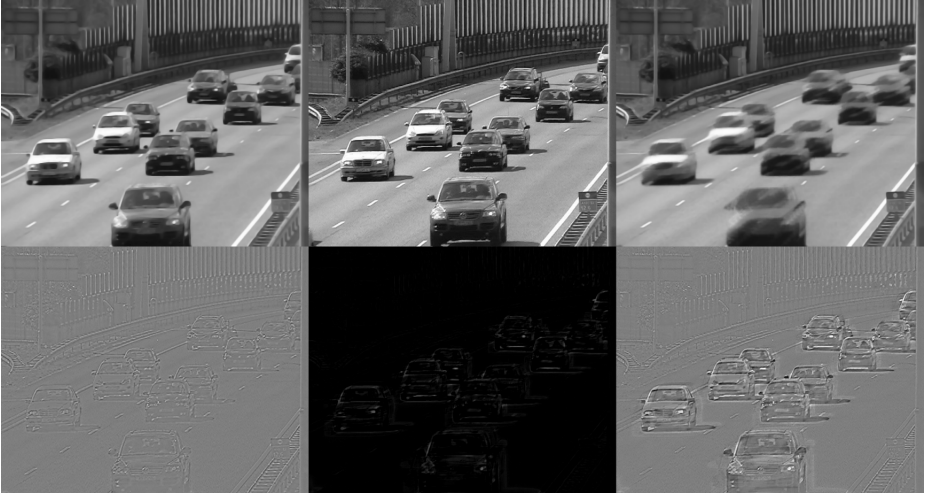


Fig. 2. From left to right, the geometrical component, U , in classic color decomposition (top) and its texture and noise component V (bottom). The original image (center and top), the difference between static and dynamic V component (center and bottom). Then our new dynamic decomposition components (right).

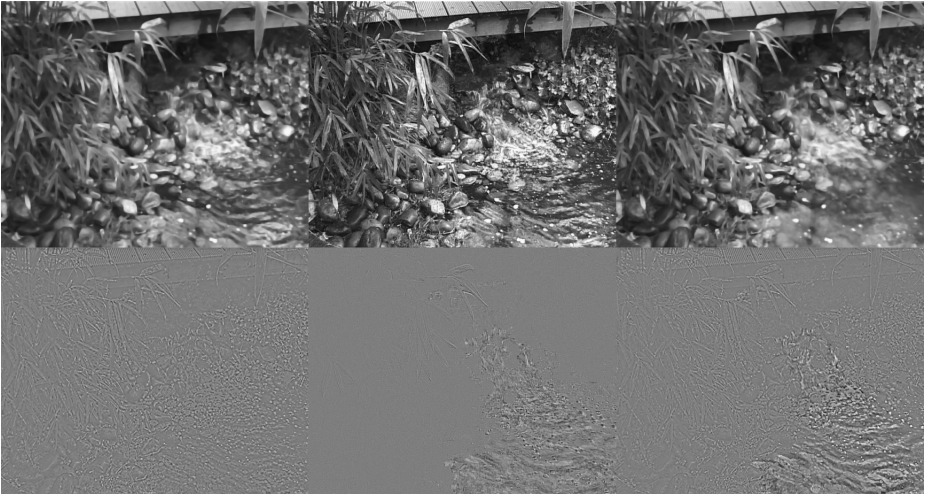


Fig. 3. From left to right, the geometrical component, U , in classic color decomposition (top) and its texture and noise component V (bottom). The original image (center and top), the time influence g_3 part of V component (center and bottom). Then our new dynamic decomposition components (right).

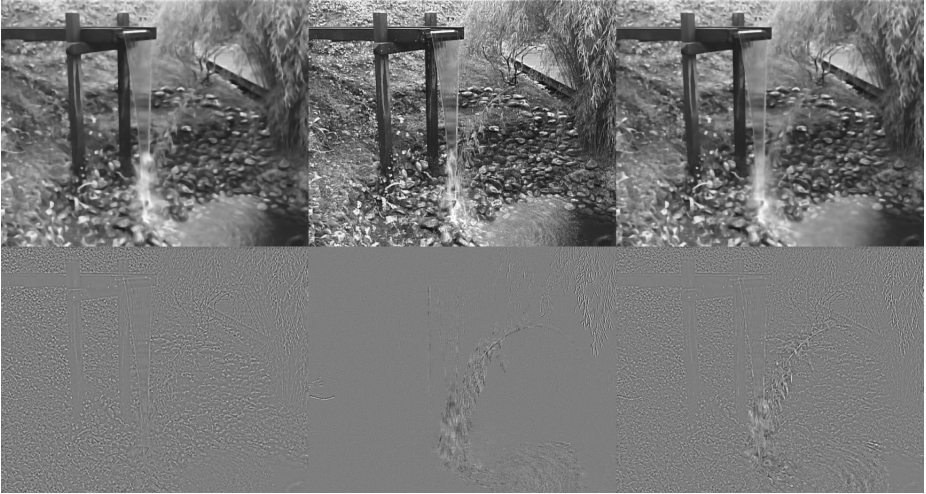


Fig. 4. From left to right, the geometrical component, U , in classic color decomposition (top) and its texture and noise component V (bottom). The original image (center and top), the time influence g_3 part of V component (center and bottom). Then our new dynamic decomposition components (right).

methods (static and dynamic decomposition are both computed with same classic parameters). We can easily see that time impact in result, water in Fig.1 is well regularized and fluid aspect is well represent in the V component. In Fig. 2 we can clearly see the reenforcement of moving cars texture without that static part and objects present in sequence are taken into account. Moreover the regularization is more robust to illumination and movement constrain.

Moreover, if user tunes parameters to obtain stronger regularization, our algorithm is able to catch wider waves in spatiotemporal texture component: see the circumference of fountain in Fig. 1, more regularized (in U component) than wider waves. It's a matter of deep in spatiotemporal texture extraction, which our algorithm is able to deal with.

4.2 Time Impact in Our Decomposition

We present, in Fig. 3, a part of a decomposed sequence of flowing water under wood bridge. We can see the static aspect of U component, regularized in space and in time, seems to be frozen, although texture component, V , present a real dynamic, strengthened by time influence. Only moving things or objects presenting dynamicity are taken in account in the third part of V component (i.e g_3). In Fig. 4 we can distinctly see time influence, reed's branch oscillating under water flow is clearly highlight, waves present in basin's fountain are well regularized in U component, water dynamicity is totally catch as texture. We can clearly see that is due to the third part of texture component, movement information is well captured in this component. In this way we obtain the dynamicity present

in video through oscillations along time dimension. These results will be useful for future work on dynamic texture characterization.

References

1. Starck, J.L., Elad, M., Donoho, D.L.: Image decomposition via the combination of sparse representations and a variational approach. *IEEE Trans. Image Processing* 14(10), 1570–1582 (2005)
2. Aujol, J.F., Aubert, G., Feraud, L.B., Chambolle, A.: Image decomposition into a bounded variation component and an oscillating component. *Journal of Mathematical Imaging and Vision* 22(1), 71–88 (2005)
3. Aujol, J.F., Gilboa, G., Chan, T., Osher, S.: Structure-texture image decomposition - modeling, algorithms, and parameter selection. *International Journal of Computer Vision* 67(1), 111–136 (2006)
4. Aujol, J.F., Chambolle, A.: Dual norms and image decomposition models. *International Journal of Computer Vision* 63(1), 85–104 (2005)
5. Aujol, J.F., Kang, S.H.: Color image decomposition and restoration. *J. Visual Communication and Image Representation* 17(4), 916–928 (2006)
6. Vese, L.A., Osher, S.J.: Modeling textures with total variation minimization and oscillating patterns in image processing. *J. Sci. Computing* 19, 553–572 (2003)
7. Vese, L.A., Osher, S.J.: Image denoising and decomposition with total variation minimization and oscillatory functions. *Journal of Mathematical Imaging and Vision* 20(1-2), 7–18 (2004)
8. Vese, L.A., Osher, S.: Color texture modeling and color image decomposition in a variational-PDE approach. In: SYNASC, pp. 103–110. IEEE Computer Society, Los Alamitos (2006)
9. Meyer, Y.: Oscillating patterns in image processing and nonlinear evolution equations. *University Lecture Series*, vol. 22, American Mathematical Society, Providence, RI, The fifteenth Dean Jacqueline B. Lewis memorial lectures (2001)
10. Ambrosio, L., Fusco, N., Pallara, D.: Functions of bounded variation and free discontinuity problems. *Oxford Mathematical Monographs*. Oxford Mathematical Monographs. Clarendon Press, Oxford (2000)
11. Aubert, G., Kornprobst, P.: *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, 2nd edn. Applied Mathematical Sciences, vol. 147. Springer, Heidelberg (2006)
12. Péteri, R., Huskies, M., Fazekas, S.: Dyntex: A comprehensive database of dynamic textures (2005), <http://www.cwi.nl/projects/dyntex/>
13. Lugiez, M., Ménard, M., El-Hamidi, A.: Spatiotemporal extension of color decomposition model and dynamic color structure-texture extraction (2008)