

Decomposition of dynamic textures using Morphological Component Analysis: a new adaptative strategy

Sloven Dubois
*Laboratoire Mathématiques
Image et Applications
La Rochelle, France
sloven.dubois01@univ-lr.fr*

Renaud Péteri
*Laboratoire Mathématiques
Image et Applications
La Rochelle, France
renaud.peteri@univ-lr.fr*

Michel Ménard
*Laboratoire Informatique
Image et Interaction
La Rochelle, France
michel.menard@univ-lr.fr*

Abstract

The research context of this work is dynamic texture analysis and characterization. Many dynamic textures can be modeled as a large scale propagating wave and local oscillating phenomena. The Morphological Component Analysis algorithm (MCA) is used to retrieve these components using a well chosen dictionary. We define a new strategy for adaptive thresholding in the MCA framework, which greatly reduces the computation time when applied on videos. Tests on synthetic and real image sequences illustrate the efficiency of the proposed method and future prospects are finally exposed.

1 Introduction

The study of Dynamic Textures (DT) is a recent research topic in the field of video processing. A DT can be described as a time varying phenomenon with a certain repetitiveness in both space and time. A flag in the wind, ripples at the surface of water, smoke or an escalator are all examples of DT. Rather than a simple extension of static textures to the time domain, a DT is a more complex phenomenon resulting from several dynamics. Their study is an active research topic with many applications such as synthesis [9], segmentation [5] or characterization [14]. The context of our work is the indexation of DT for automatic video retrieval [7].

Each DT has its own characteristics, such as stationary, repetitiveness, velocity ... On Figure 1 showing an image sequence of a river surface, two motions can be observed: a high frequency motion

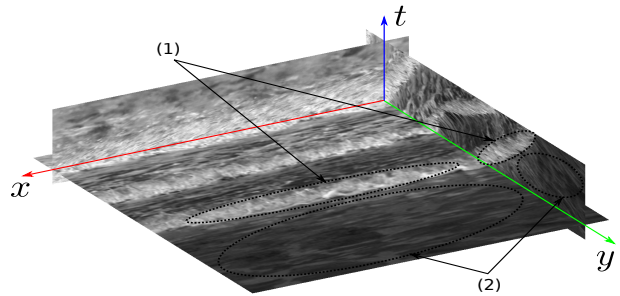


Figure 1. 2D+T slices of a dynamic texture: local oscillating phenomenon (2) and long range propagating wave (1).

(2) carried by an overall motion (1). Many DT can be decomposed into one or several local oscillating motions carried by far range waves. In order to better characterize these two sets of components, it is necessary to extract them separately.

In this article, the Morphological Component Analysis (MCA) is used for decomposing and analyzing image sequences of natural scenes. To our knowledge, the only existing work using MCA and video is recent and focuses on the inpainting of a cartoon sequence [13].

In Section 2, the MCA is briefly described. The dictionaries selected in the MCA, adapted to the model used for DT, are then presented. A key issue is the computation time of MCA that is related to the thresholding strategy. We propose in Section 3 a new adaptive thresholding strategy that reduced the computation time by a factor of four compared to the original algorithm. Results on synthetic and real sequences of DT are presented and future prospects are discussed in Section 4.

2 Decomposing a DT using MCA

According to researches on synthesis [10] and observations made on a large DT database [11], a DT can be modelled as a sum of local oscillations carried by longer range waves. Recent works for decomposing images and videos [12, 4, 1] seem relevant for extracting these components. We have chosen the MCA because of the richness of the available dictionary, which is crucial regarding the complexity of DT.

Given a signal y described as a linear superposition of morphological components K disturbed by a noise ε : $y = \sum_{i=1}^K y_i + \varepsilon$.

The MCA approach allows to find an acceptable solution to the inverse problem of decomposing a signal on a given vectorial basis, *i.e.* to extract components $(y_i)_{i=1,\dots,K}$ from a degraded observation y according to a sparsity constraint. The MCA approach assumes that each component y_i can be represented sparsely in the associated basis Φ_i : $\forall i = 1, \dots, K, y_i = \Phi_i \alpha_i$.

Algorithm 1 describes the main steps of the MCA. A detailed description can be found in [12].

Algorithm 1 Morphological Component Analysis

```

1. while  $\left\| y - \sum_{j=1}^K \tilde{y}_j^{(k-1)} \right\|_2 \leq \sigma$  do
2.   // For each component
3.   for  $i = 1$  à  $K$  do
4.     // Compute the marginal residual
5.      $\tilde{r}_i^{(k)} = y - \sum_{j=1}^{i-1} \tilde{y}_j^{(k)} - \sum_{j=i+1}^K \tilde{y}_j^{(k-1)}$ 
6.     // Projection of  $\tilde{r}_i^{(k)}$  on basis  $\Phi_i$ 
7.      $\alpha_i^{(k)} = \Phi_i^T \left( \tilde{r}_i^{(k)} \right)$ 
8.     // The new estimation of  $y_i$ 
9.      $\tilde{y}_i^{(k)} = \Phi_i \left( \delta_{\lambda^{(k)}} \alpha_i^{(k)} \right)$ 
10.  end for
11.  // Update of the threshold  $\lambda$ 
12.   $\lambda^{(k+1)} = \text{update}(\lambda^{(k)}, \text{strategy})$ 
13. end while
```

A crucial point in the MCA approach is the dictionary definition. An unsuitable choice of transformations will lead to non sparse and irrelevant decompositions of the different dynamical phenomena present in the sequence. As mentioned previously, we model DT as a sum of local oscillations carried by long range waves. It is therefore necessary to associate each component with the most representative basis. In [6], the authors show that the curvelet transform [3] is relevant

for extracting non-local phenomena propagating temporally. It thus seems particularly interesting to model long range waves present in a DT.

The second part of the model is composed of locally oscillating phenomena that will be extracted using the local cosine transform.

The dictionary that we use in the MCA algorithm is then composed of the curvelet transform Φ_1 and the local cosine transform Φ_2 .

3 Thresholding strategy

Context

The computation time of decomposition algorithms represents a major challenge for sequence analysis (indexing and browsing). Some transforms require several minutes of computation on a short image sequence. A recent work [2] has shown that a hundred of iterations is necessary to establish a good separation of the different components when a linear thresholding strategy (LTS) is used. In our case, the total computational time for a 5 second sequence is given by: $100 * [T(\Phi_1^T) + T(\Phi_1) + T(\Phi_2^T) + T(\Phi_2)] = 20 \text{ hours}$, where $T()$ measures the execution time of one projection on Φ_i during one cycle of the algorithm.

If we extend this result to the entire DT database DynTex [11], 583 days of calculation are required. Recently Bobin *et al.* [2] have proposed a thresholding strategy 'Mean of Max' (MOMS) that enables to obtain similar results but with fewer iterations (25 in average instead of 100). It represents a computation time of approximately 7 hours 30 for a 5 second video, resulting in approximately 219 days for the whole database.

For indexing the DynTex database, the computation time of the MOMS is still acceptable, since it is always possible to divide the workload on several processors. In the case where one searches for a particular texture using a query sequence, these calculations are acceptable only on sequences of limited duration and of low resolution. One goal of this work is to reduce these limitations by proposing a new thresholding strategy.

An adaptive thresholding strategy

The quality of the results from the decomposition of a signal using the MCA algorithm strongly depends on the evolution of the threshold $\lambda^{(k)}$ in one iteration (in one *for* loop). Figure 2 shows two different evolutions of $\lambda^{(k)}$ corresponding to two strategies (1) and (2). The evolution of $\lambda^{(k)}$

is slower in case (1) than in (2). In this example, evolution (1), respectively (2), leads to select 5%, respectively 25%, in the two bases. If we consider that evolution (1) gives an optimal threshold here, a failure to control the value of $\lambda^{(k)}$ (in case 2) will lead to a rapid allocation of too many coefficients in the two bases, degrading the final decomposition.

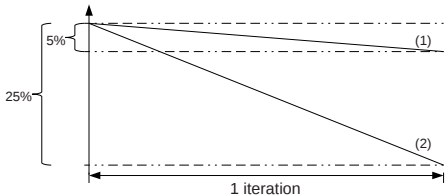


Figure 2. Two thresholding strategies leading to different evolutions of the threshold value (in one *for* loop).

The linear thresholding strategy (LTS) leads to the optimum $\lambda^{(k)}$ for 100 iterations [2]. In a large number of natural textures, this number of iterations can be greatly reduced, depending on the texture itself. LTS is then no longer optimum. However, the threshold evolution using LTS can be considered as a minimum slope below which the evolution of $\lambda^{(k)}$ is sub-optimal. A good strategy for the calculation of $\lambda^{(k)}$ must lead to a slope greater than or equal to the one obtained using LTS. The 'Mean of Max' strategy (MOMS) is interesting as it can adaptively change the evolution of $\lambda^{(k)}$. However, on natural texture sequences, this strategy tends to reduce too drastically this slope, or even almost cancel it.

We propose to combine these two strategies into a new so-called adaptive thresholding strategy (ATS), which defines $\lambda^{(k)}$ as the minimum value of $\lambda^{(k)}$ calculated using strategies LTS and MOMS.

The $\lambda^{(k)}$ update using ATS is formalized as follows:

$$\lambda^{(k)} = \min \left(\frac{1}{2}(m_1 + m_2), \lambda^{(k)} - \frac{\lambda^{(1)} - \lambda_{\min}}{100} \right)$$

with :

$$m_1 = \max_i \|\Phi_i^T r^{(k)}\|_{\infty}, \quad m_2 = \max_{j \neq i} \|\Phi_j^T r^{(k)}\|_{\infty}$$

$$r^{(k)} = y - \sum_{j=1}^K \tilde{y}_j^{(k)}$$

is the residual term

Using this strategy, we are sure to change the value of $\lambda^{(k)}$ corresponding to the steepest slope. In other words, when MOMS leads to values of $\lambda^{(k)}$ evolving slowly, $\lambda^{(k)}$ follows the LTS $\lambda^{(k)} = \lambda^{(k)} - \frac{\lambda^{(1)} - \lambda_{\min}}{100}$. Otherwise, $\lambda^{(k)}$ follows

the MOMS, reducing the number of *for* loops in the algorithm and the computation time.

4 Applications

Figure 3.a shows a video of water generated by our DT model (not detailed here due to a lack of space, see [8]). After our MCA decomposition scheme, we are able to retrieve the long term wave (fig. 3.b) and the local oscillations (fig. 3.c) used in the synthesis. This reinforces the choice of the chosen dictionary composed of the curvelet and local cosine transforms.

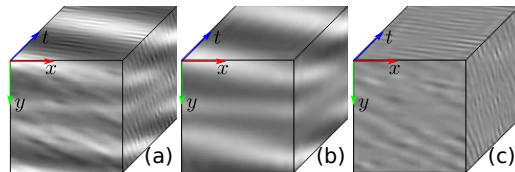


Figure 3. MCA decomposition of a synthetic video (a) on the curvelet basis (b) and the local cosine basis (c)

Applied on a real sequence from DynTex of a duck on streaming water (fig. 4.a), our algorithm is still able to separate geometrical components (fig. 4.b) from more local oscillations (fig. 4.c)¹. One can observe on the spatio-temporal (xt) slices of figure 4 and on figure 4.b, that the reflection of trees in water are better observable in the curvelet component sequence than in the original video.

Let us point out that the use of ATS enables a significant gain in computation time. Indeed, in average for one video sequence, the computation time is about 4 hours 30 (instead of 7h30), which leads to 131 days for the whole DynTex database.

5 Conclusion and prospects

This paper deals with the decomposition of DT in videos into different dynamical components. We show that the MCA algorithm is well suited for this decomposition in a well chosen basis, but suffers from significant computation time. We propose a new thresholding strategy which leads to a significant gain in the algorithm speed.

Other thresholding strategies are being studied to further improve this computation time. It is particularly necessary to develop strategies that take

¹This video and other results are visible at: http://mia.univ-larochelle.fr/demos/dynamic_textures/

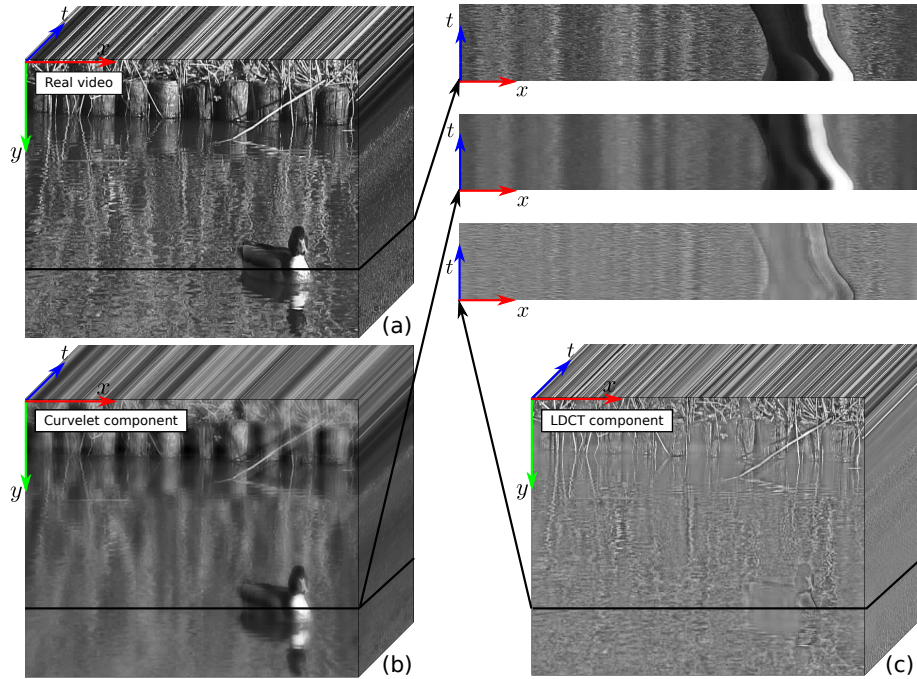


Figure 4. MCA decomposition of a real video (a) on the curvelet basis (b) and the local cosine basis (c)

better into account our proposed model and the features of natural DT. The extracted components of DT can later be used as features for video retrieval applications.

References

- [1] J.-F. Aujol and A. Chambolle. Dual norms and image decomposition models. *Int. J. Comput. Vision*, 63:85–104, 2005.
- [2] J. Bobin, J.-L. Starck, J. Fadili, Y. Moudden, and D. Donoho. Morphological component analysis : An adaptive thresholding strategy. *IEEE Trans. on Image Processing*, 16:2675–2681, 2007.
- [3] E. Candès, L. Demanet, D. Donoho, and L. Ying. Fast discrete curvelet transforms. *Multiscale Modeling & Simulation*, 5:861–899, 2006.
- [4] T. Chan, S. Osher, and J. Shen. The digital TV filter and nonlinear denoising. *IEEE Trans. on Image Processing*, 10:231–241, 2001.
- [5] G. Doretto, D. Cremers, P. Favaro, and S. Soatto. Dynamic texture segmentation. In *ICCV'03*, pages 1236–1242, 2003.
- [6] S. Dubois, R. Péteri, and M. Ménard. A 3D discrete curvelet based method for segmenting dynamic textures. In *ICIP'09*, pages 1373–1376, Cairo, Egypt, 2009.
- [7] S. Dubois, R. Péteri, and M. Ménard. A comparison of wavelet based spatio-temporal decomposition methods for dynamic texture recognition. In *LNCS, IbPRIA'09*, volume 5524, pages 314–321, Povoia de Varzim, Portugal, 2009.
- [8] S. Dubois, R. Péteri, and M. Ménard. Modèle de Textures Dynamiques et décomposition par l'approche MCA. Internal Report of MIA and L3i laboratories. <http://hal.archives-ouvertes.fr/hal-00449634/en/>, 2010.
- [9] J. Filip, M. Haindl, and D. Chetverikov. Fast synthesis of dynamic colour textures. In *ICPR'06*, pages 25–28, Hong Kong, 2006.
- [10] M. Finch. *GPU Gems: Programming Techniques, Tips, and Tricks for Real-Time Graphics, Chap.1*. Randima Fernando.
- [11] R. Péteri, S. Fazekas, and M. J. Huiskes. DynTex: a comprehensive database of Dynamic Textures. *Pattern Recognition Letters*, 2010. <http://projects.cwi.nl/dyntex/>. To appear.
- [12] J.-L. Starck, M. Elad, and D. Donoho. Image decomposition via the combination of sparse representations and a variational approach. *IEEE Trans. on Image Processing*, 14:1570–1582, 2004.
- [13] A. Woiselle, J.-L. Starck, and J. Fadili. Inpainting with 3D sparse transforms. In *GRETSI'09*, Dijon, France, 2009.
- [14] G. Zhao and M. Pietikainen. Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE Trans. on PAMI*, (29):915–928, 2007.