Information Sciences 328 (2016) 510–527

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Group-based image decomposition using 3-D cartoon and texture priors

Tian-Hui Ma, Ting-Zhu Huang*, Xi-Le Zhao

School of Mathematical Sciences/Research Center for Image and Vision Computing, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China

ARTICLE INFO

Article history:
Received 29 March 2015
Revised 6 August 2015
Accepted 23 August 2015
Available online 10 September 2015

Keywords:
Cartoon-texture decomposition
Convex optimization
Low-rank
Nonlocal self-similarity
Parameter selection

ABSTRACT

We propose a novel image decomposition method to decompose an image into its cartoon and texture components. To exploit the nonlocal self-similarity of cartoon-plus-texture images, we construct groups by stacking together similar image patches into 3-D arrays and consider group as the basic unit of decomposition. We decompose each group via a convex optimization model consisting of 3-D cartoon and texture priors. These priors characterize the local properties of the cartoon and texture components and the nonlocal similarity within each component in a unified and natural manner. We develop the alternating direction method of multipliers (ADMM) to efficiently solve the proposed model. For further improvement, we investigate an adaptive rule for the estimation of the regularization parameter. The proposed method is also extended to tackle noisy images. Numerical experiments confirm that the performance of the proposed method is competitive with some of the state-of-the-art schemes.

1. Introduction

Image decomposition is a fundamental problem in image processing and computer vision. It plays a significant role in many real-world applications such as astronomical imaging [19], image restoration [3], and image analysis [39]. The objective of image decomposition is to decompose the input image into the sum of two meaningful components: the cartoon component and the texture component. The cartoon component contains the global geometrical configurations of the image, such as contours, homogeneous regions, and sharp edges, while the texture component includes the high-frequency information, such as fine structures and repeating patterns. Image decomposition is a typical inverse problem since a multiplicity of possible solutions exists. To obtain meaningful results, the regularization techniques, which stabilize inverse problems by introducing some priors, have been increasingly applied to image decomposition over the past decades [1,3,7,16,19,34,36,38,40,41,44,46,48,49].

The key issue in regularization techniques is the design of the priors, which encourage the solution to exhibit some expected properties. For the cartoon component, the most frequently chosen prior is the total variation (TV) [44], because of its advantage of restoring piecewise constant functions while preserving edges. TV has been studied for decades; see [18,30,37,53] for some recent developments. For the texture component, its characterization is, however, still a challenging difficulty due to the great variety of texture. Over the years, various research efforts have been devoted to this topic and many texture priors have been developed. One successful proposal is the G-norm [36], which effectively characterizes the functions with large oscillations. The G-norm has been extensively studied with theoretical properties [19], numerical implementations [1,41,49], and applications

* Corresponding author. Tel.: +86 28 61831608; fax: +86 28 61831280.
E-mail addresses: nkmth0307@126.com (T.-H. Ma), tingzhuhuang@126.com, tzhuang@uestc.edu.cn (T.-Z. Huang), xlzhao122003@163.com (X.-L. Zhao).

http://dx.doi.org/10.1016/j.ins.2015.08.039
0020-0255/

[3,19,38]. Recently, the G-norm has been applied to the problem of coupled image decomposition and restoration [38,52]. However, the G-norm fails to differentiate texture from noise, as both texture and noise have an oscillating nature, which limits its application to noisy images. Another texture priors assume that texture has a sparse representation in suitable transform domains or over redundant dictionaries. Hence, the texture component is extracted by choosing a suitable representation basis and by promoting the sparsity of the representation coefficients. Examples of such methods can be found in [7,34,48], and the reader is also referred to the work [16] for a detailed overview of the literature on this topic. However, if the sparsity assumption does not hold, those methods tend to erroneously extract some cartoon-like features and to compromise small magnitude texture.

Lately, low-rank texture priors have been developed to capture well-patterned texture by exploiting the low-rankness within the texture component. In [46] the authors have proposed a low-rank texture prior called the low patch-rank (LPR). By promoting the linear dependence of texture patches, LPR effectively extracts well-patterned features. However, the main drawback of LPR comes from its global treatment of texture as described in [40], i.e., texture is considered as a global and well-patterned structure with only a few individual patterns. This makes LPR perform poorly on images with various different texture patterns and create undesirable pattern-like artifacts, especially when processing noisy images. Very recently, a low-rank texture prior has been proposed using the block nuclear norm (BNN) [40]. Different from treating texture as a whole, the BNN prior considers texture as a globally dissimilar but locally well-patterned structure. Motivated by this intuitive observation, the BNN model encourages the low-rankness of local texture patches with possible overlap and shear. It has been shown that BNN overcomes the drawback of LPR and produces higher quality decomposition results.

In recent years, the nonlocal self-similarity has emerged as one of the most significant properties of natural images, which depicts the redundancy of local image features that are similar to each other. Compared with the conventional image assumptions mentioned above, the nonlocal self-similarity prior is highly adaptive and provides a more effective characterization of natural images. Due to its potential, the nonlocal methods have drawn increasingly more research attention in the past decade. Inspired by the success of the NL-means denoising filter [5], a number of advanced nonlocal methods have been developed in various image processing applications, such as denoising [10,42,56], restoration [13,25,57], inpainting [11], segmentation [18], compressive sensing [12], and super-resolution [23,28]. However, to the best of our knowledge, the effectiveness of the nonlocal self-similarity prior to image decomposition has not been documented in the open literature.

In this paper we investigate a novel image decomposition method combining the regularization techniques and nonlocal methods, aiming at exploiting both the local properties and the nonlocal self-similarity of cartoon-plus-texture images. For this purpose we propose to decompose the input image in the unit of group, i.e., a 3-D array consisting of similar image patches. The proposed method consists of constructing groups by stacking together similar image patches, decomposing each group into its cartoon and texture components, and aggregating the obtained estimations to form the final result. Such group-based decomposition scheme has the following advantages:

- a group containing only a few individual patterns is easier to deal with than the whole image having various different patterns;
- the similarity between the grouped patches also implies the similarity between the underlying cartoon and texture estimates, and decomposition can significantly benefit from an efficient exploitation of this similarity;
- group-based decomposition has a redundancy nature which helps to avoid typical artifacts and makes our method effective and robust for noisy images.

In order to decompose each group, we design a convex optimization model involving 3-D cartoon and texture priors. Specifically, the cartoon component is characterized by the 3-D total variation (3-D TV) [9] and the texture component is extracted by the tensor trace norm (TTN) [31], which are, respectively, generalizations of the 2-D priors TV and matrix trace norm to the 3-D cases. As we will show in Section 3.2, these priors characterize both the local properties and the nonlocal self-similarity within each component. We develop the alternating direction method of multipliers (ADMM) [17,20,37] to efficiently solve the proposed model. For further improvements, we investigate an adaptive rule for the estimation of the regularization parameter, in order to control the trade-off between the two priors. Moreover, the proposed method is extended to handle noisy images. Experimental results have shown that the performance of the proposed method is competitive with the state-of-the-art schemes LPR [46], BNN [40], and ADMGB [38].

In summary, our work has the following two main contributions:

- We adopt group as the basic unit of decomposition, providing an efficient exploitation of the nonlocal self-similarity of cartoon-plus-texture images.
- We propose a convex optimization model involving 3-D cartoon and texture priors to efficiently decompose each group.

The outline of the rest of the paper is as follows. In the next section, we introduce some notations. Section 3 is devoted to a detailed description of our method, and an extension to the noisy image scenario can be found in Section 4. Section 5 reports the results of our numerical experiments. Finally, we conclude the paper in Section 6.

2. Notations

From now on, we will restrict our attention to the discrete setting. We use boldface Euler script letters for tensors, e.g., $\boldsymbol{\mathcal{A}}$, boldface capital letters for matrices, e.g., $\mathbf{A}$, boldface lowercase letters for vectors, e.g., $\mathbf{a}$, and lowercase letters for scalars, e.g., a.