J. Vis. Commun. Image R. 38 (2016) 230–245

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Context-aware joint dictionary learning for color image demosaicking ☆,☆☆

Kai-Lung Hua [a], Shintami Chusnul Hidayati [a], Fang-Lin He [a], Chia-Po Wei [b], Yu-Chiang Frank Wang [b,*]

[a] Dept. of CSIE, National Taiwan University of Science and Technology, Taipei, Taiwan
[b] Research Center for Information Technology Innovation, Academia Sinica, Taipei, Taiwan

ARTICLE INFO

Article history:
Received 17 October 2014
Revised 23 October 2015
Accepted 1 March 2016
Available online 9 March 2016

Keywords:
Color demosaicking
Dictionary learning
Self-learning
Sparse representation

ABSTRACT

Most digital cameras are overlaid with color filter arrays (CFA) on their electronic sensors, and thus only one particular color value would be captured at every pixel location. When producing the output image, one needs to recover the full color image from such incomplete color samples, and this process is known as demosaicking. In this paper, we propose a novel context-constrained demosaicking algorithm via sparse-representation based joint dictionary learning. Given a single mosaicked image with incomplete color samples, we perform color and texture constrained image segmentation and learn a dictionary with different context categories. A joint sparse representation is employed on different image components for predicting the missing color information in the resulting high-resolution image. During the dictionary learning and sparse coding processes, we advocate a locality constraint in our algorithm, which allows us to locate most relevant image data and thus achieve improved demosaicking performance. Experimental results show that the proposed method outperforms several existing or state-of-the-art techniques in terms of both subjective and objective evaluations.

1. Introduction

A digital color image is made up of pixels, which consist of three independent primary color components: red, green, and blue. In order to reduce costs like hardware size, cost, or power consumption, most digital cameras capture one color component at each pixel location by a single monochromatic image sensor (CCD or CMOS), which is overlaid with a color filter array (CFA). As shown in Fig. 1, the most commonly used CFA is the Bayer pattern [1], which places green color pixels in a quincunx lattice and red/blue color pixels in rectangular ones. To obtain the full resolution color image from such CFA samples, two unknown color components need to be estimated from neighboring pixels. Since the CFA samples carry mosaic patterns for different colors, this color restoration process has been widely known as *demosaicking*.

To address the problem of demosaicking problems and consequently many approaches have been proposed, interpolation-based methods such as bilinear, bicubic, and spline interpolation were initially proposed. However, this type of approaches yield severe artifacts like false color information or zippering, especially along the edges or highly textural regions due to the smooth transition. To reduce such artifacts, recent works suggest edge-directed interpolation and/or the joint exploitation of both intra and inter-channel dependencies during the interpolation process for demosaicking [2–9].

Nevertheless, due to the abundance and diversity of natural images, it is still very challenging to solve the image demosaicking problem through direct analytical analysis or construction of image processing models. Lately, researchers further advance machine learning techniques for tackling this problem. For example, Zhang et al. [10] considered the calculation of intra and inter-band interpolation, and applied minimum mean-square error (LMMSE) and support vector regression (SVR) to determine the weighting factors for the above interpolation for producing the final demosaicked output. Mairal et al. [11,12] proposed non-local sparse models for image restoration. Their proposed method is based on the idea of jointly decomposing groups of similar signals on subsets of the learned dictionary. Wu et al. [13] proposed a sparsity-based method by introducing the sparse representations for both intra and inter-channels. They derived such representations via an ℓ_1 minimization formulation, and thus the observed model can be applied for demosaicking.

☆ This paper has been recommended for acceptance by O. Au.

☆☆ This work was supported in part by Ministry of Science and Technology of Taiwan via MOST104-2221-E-011-091-MY2, MOST103-2221-E-011-105, and MOST103-2221-E-001-021-MY2.

* Corresponding author.
E-mail address: ycwang@citi.sinica.edu.tw (Y.-C.F Wang).

http://dx.doi.org/10.1016/j.jvcir.2016.03.004
1047-3203/

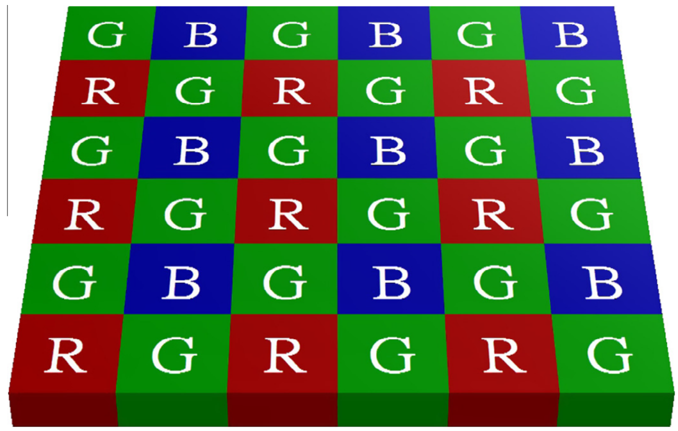

Fig. 1. Bayer pattern of color mosaic for digital cameras. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

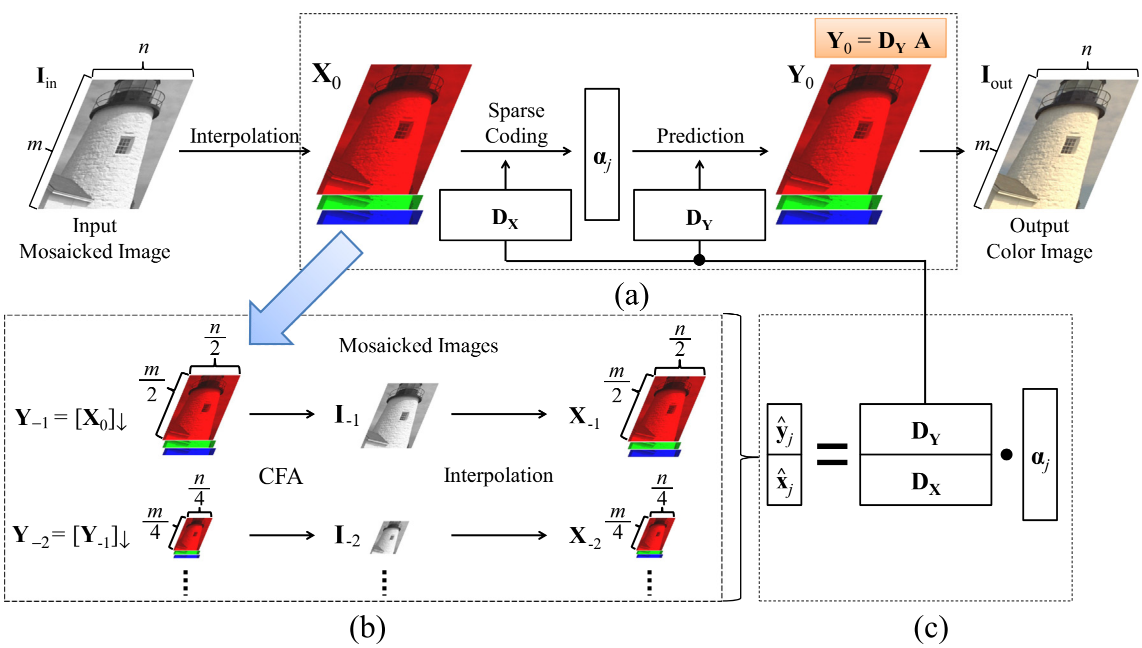

Fig. 2. Flowchart of our proposed self-learning demosaicking approach via locality-sensitive joint dictionary learning. (a) The demosaicking process using the joint dictionary observed during the self-learning process. (b) Producing self-learning data of ground truth images and their mosaicked versions from the input mosaicked image. (c) Learning of locality-sensitive joint dictionary for associating data observed in (b).

It is worth noting that, the common disadvantage of most prior learning-based approaches is the requirement of training data collection in advance. Without the presence of training data, one cannot observe the associated learning models for demosaicking. Inspired by recent works on single image super-resolution [14–16], we propose a self-learning color demosaicking algorithm in this paper. Our approach allows one to produce self-training data of interest directly from the input mosaicked image. Based on the success of sparse representation for image processing tasks, we advocate the learning of *locality-sensitive joint dictionary* for describing the mosaicked images and their ground truth versions during the training (self-learning) stage. Besides, considering the color and textural context information, we learn dictionaries for different context categories. Once the dictionary is observed, we can apply the dictionary atoms[1] on the input mosaicked image to synthesize/recover the demosaicked output directly. As a result, there is no need to collect external training image data beforehand, which makes the proposed learning-based demosaicking method much more practical (and preferable for real-world applications). We compare the proposed demosaicking algorithm with several state-of-the-art approaches [17,18,6,19] for several images. Experimental results show that we are able to achieve comparable or improved performance in terms of both subjective and objective quality measures, while *no* training data collection in advance is required for our method.

The rest of the paper is organized as follows. Section 2 reviews the sparse representation. Our proposed locality-sensitive joint dictionary learning algorithm is presented in Section 3, and experimental results are presented in Section 4. Finally, Section 5 summarizes this paper.

[1] Atoms are defined as a set of basis vectors in the dictionary.