Engineering Science and Technology, an International Journal 19 (2016) 1067–1075

HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

Engineering Science and Technology, an International Journal

journal homepage: http://www.elsevier.com/locate/jestech

Full Length Article

# Shadow removal using sparse representation over local dictionaries

Remya K. Sasi [a,*], V.K. Govindan [b]

[a] NIT Calicut, Kerala, India
[b] IIIT, Kerala, India

ARTICLE INFO

*Article history:*
Received 24 September 2015
Received in revised form
29 November 2015
Accepted 1 January 2016
Available online 19 February 2016

*Keywords:*
Shadow
Dictionary learning
Shadow removal
Sparse coding
MCA

ABSTRACT

The presence of shadow in an image is a major problem associated with various visual processing applications such as object recognition, traffic surveillance and segmentation. In this paper, we introduce a method to remove the shadow from a real image using the morphological diversities of shadows and sparse representation. The proposed approach first generates an invariant image and further processing is applied to the invariant image. Here, shadow removal is formulated as a decomposition problem that uses separate local dictionaries for shadow and nonshadow parts, without using single global or fixed generic dictionary. These local dictionaries are constructed from the patches extracted from the residual of the image obtained after invariant image formation. Finally, non-iterative Morphological Component Analysis-based image decomposition using local dictionaries is performed to add the geometric component to the non-shadow part of the image so as to obtain shadow free version of the input image. The proposed approach of shadow removal works well for indoor and outdoor images, and the performance has been compared with previous methods and found to be better in terms of RMSE.



## 1. Introduction

Shadow detection and removal process is widely used as a preprocessing operation in various image processing applications for the removal of undesirable noise and objects. For example, applications such as video surveillance [1], scene interpretation [2] and object recognition [3] require shadow removal as an initial step to eliminate the undesirable effect of noise on the performance of such systems. Once detected, shadows in images are used for applications such as detection of object shape and size in aerial images, detection of movement of objects in video surveillance system, and finding the number of light sources and illumination conditions in natural images. In digital photography, removal of shadows can help to improve the visual quality of photographs. Ignoring the existence of shadows in images can, in general, degrade the output quality.

Shadow detection and removal is an active research area for the last two decades. Several algorithms have been proposed based on learning [4], color models [5], region [6,7] and invariant image models [8,9] for shadow detection and removal in image as well as in videos. A major work by Lalonde et al. [10] mainly focuses on the shadows cast by objects onto the ground plane. Other notable works are based on assumptions of Lambertian reflectance and Planckian lighting [11]. Interested readers can see a review article by Sasi and Govindan [12] to get a more comprehensive report of the methodologies reported in the field of shadow detection and removal during the last decade.

Though shadow removal involving multiple images [13] and interactive methodologies [14–16] provides superior performance, fully automatic approaches available for single image shadow removal stand behind them in terms of performance. This is because of the fact that indoor and outdoor shadows are much affected by the direction, intensity of light source, as well as geometry and texture of the objects where shadow is cast.

The review carried out reveals that the research work reported in shadow detection and removal works satisfactorily in the case of user interaction [14–16] and for multiple images [13]. The automatic approaches available for single shadow removal are more complex to implement and set much restriction in the class of images under consideration [10,11]. Reintegration methods and local area processing are time intensive. Also, many methods cannot distinguish between near dark objects and shadows. In the case of small-patch regions, image in-painting method is more suitable, whereas in-painting large patch holes involve huge computation. Thus, the topic of single image shadow detection and removal requires a good amount of further research to develop an approach that provides satisfactory performances.

This paper proposes a method to remove shadow from a real image using sparse representation and a variation of MCA. Sparsity is a powerful way to approximate signals and images by using a sparse

* Corresponding author. Tel.: +919496812035.
*E-mail address:* remyaksasi@gmail.com (R.K. Sasi).
Peer review under responsibility of Karabuk University.

http://dx.doi.org/10.1016/j.jestch.2016.01.001
2215-0986/

linear combination of atoms from an over-complete dictionary. Sparse representation is being used in signal and image processing applications such as de-noising [17], super-resolution [18], in-painting [19], deblurring [20], segmentation [21], and compression [22], demonstrating that sparse models are well suited to natural images as well. Starck et al. developed the idea of MCA in a series of papers and it was used in separating the texture from the geometric component [23–25]. MCA algorithm works by decomposing the image into edges and textures, using the morphological differences in these features [23,24]. Each of the shape features is related to a fixed dictionary of atoms such as wavelets or Discrete Cosine Transforms. Typical MCA is an iterative thresholding scheme in which the threshold linearly decreases to zero. Another similar work in the area is "Learning the Morphological Diversity" by Peyré et al. [26]. They introduced an adaptive MCA scheme by learning the morphologies of image layers. A combination of adaptive local dictionaries and fixed generic dictionaries using wavelets and curvelets is used for decomposition. The main deficiency of these models is the existence of similar atoms corresponding to cartoon and texture dictionaries that produce coherence. So, to get the expected solution, proper manual initialization is essential.

Apart from sparse representation, we are also using invariant image formation as a base step in the proposed shadow removal method. Many of the works in the area of shadow removal using invariant image formation were authored by Finlayson and his students [8,9,27–31]. In general, his methods are based on forming an invariant image, in which shadows do not appear followed by reconstructing the required missing components using reintegration. Invariant image formation results in the loss of photo quality of image. To bring back the fine details lost in the invariant image formation, reintegration is performed using Poisson equation by averaging over retinex paths [8] or Hamiltonian paths [27]. In most of the methods missing information after shadow removal is interpolated using image inpainting methods. Finlayson also limits his work to images that follow Lambertian model where Planckian illumination lights the scenes. However, real scenes need not satisfy Lambertian assumption.

The remaining part of this research report is put forth as in the following: The basics of sparse coding, dictionary learning, MCA and invariant image formation required for better readability of the paper is presented in Section 2. The proposed method of shadow removal using sparse representation is discussed in Section 3. The dataset used, results obtained and further discussions of the proposed work are given in Section 4. The paper is concluded in Section 5 highlighting the approach used and performance gain achieved.

## 2. Preliminaries

This section briefly reviews the theory behind sparse coding (SC), dictionary learning, morphological component analysis (MCA) and intrinsic image formation to better understand the proposed shadow removal methodology using MCA.

### 2.1. Sparse coding

Using sparse coding, an image $y$ can be expressed as a set of few elementary signals taken from an over-complete dictionary $A$, subject to $\alpha$ should be sparse.

$$\min_{\alpha} \|\alpha\|_0 \quad \textit{subject to} \quad y = A\alpha \tag{1}$$

Considering noise and sparsity constraint, we can add a regularization parameter $\lambda$ and reformulate (1) as

$$\min_{\alpha} \|y - A\alpha\|_2^2 + \lambda\|\alpha\|_0 \tag{2}$$

Solution to the above problem is NP hard; however, many convex [32], non-convex [33] optimization and greedy approximation algorithms [34,35] exist in literature to deal with problems having the above formulation. Since norm 2 minimization is equivalent to norm 1, L1 regularized LS also gives solution for the above problem [32,36,37].

### 2.2. Dictionary learning

In dictionary learning, the algorithm is given samples of the form $y = A\alpha$, where $\alpha \in \Re^m$ is an unknown random sparse vector and $A$ is an unknown dictionary matrix in $\Re^{n*m}$. The goal is to learn $A$ and $\alpha$ from given $y$ such that

1. A should be over-complete ie $m > n$
2. Atoms in A are linearly dependent
3. Representation error, $E = y - A\alpha$, is minimized

Dictionary learning can be formulated as given in (3). For fixed dictionary solve system of equations subject to $\alpha$ is sparse. Then, for fixed $\alpha$, update A.

$$\arg\min_{\alpha,A} \|y - A\alpha\|_2^2 \quad s.t. \quad \|\alpha\|_0 <= k \tag{3}$$

where $k$ denotes sparsity.

Dictionaries can be fixed, global or local. Global dictionaries are built from clean patches of selected images of a database. Earlier, in the sparse coding area, focus was mostly given to fixed over-complete dictionaries, such as wavelets and discrete cosine transform [38]. These approaches are called generic since the dictionary is predefined. Local dictionaries are learned online and hence more adapted to the input image. Different methodologies exist in dictionary learning literature, starting from fixed dictionaries to online dictionaries [39]. Fixed dictionaries find sparse approximations of the set of training signals for fixed dictionary, whereas optimized dictionaries keep sparse signals fixed and build optimized dictionaries [40,41]. MOD, K-SVD [42], and online dictionary learning [39,43] are the popular algorithms in the area.

### 2.3. MCA

Morphological Component Analysis is used for the separation of the components of an image having different morphologies. MCA and Basis Pursuit are based on sparsity, but MCA is much faster and is capable of handling large data sets.

Consider an image $y$ having '$s$' morphological components, such that $y = \sum_{k=1}^{S} y_k$, where $y_k$ denotes the $k$th geometric or textural component of $y$. To decompose the image $y$ into $\{y_k\}_{k=1}^{S}$, the MCA algorithm finds the sparsest solution over the dictionaries $A_k$ such that

$$\{\alpha_1, .. \alpha_S\} = \arg\min_{\{\alpha_1,..\alpha_S\}} \sum_{k=1}^{s} \|\alpha_k\|_1 + \lambda \left\| y - \sum_{k=1}^{s} A_k\alpha_k \right\|_2^2 \tag{4}$$

where $\alpha_k$ denotes the $k$th sparse solution

MCA algorithm uses fixed generic dictionaries such as wavelets and curvelets in representing geometric component and DCT for textural components. A major step in this algorithm is the selection of dictionaries that are mutually incoherent. Further details about the MCA algorithm can be found in Reference [24].

### 2.4. Invariant/Intrinsic image formation

An invariant image is invariant to illumination, color and intensity. Illumination invariant image is invariant to illumination. Shadow is caused by illumination and hence illumination invariant image