On Sparse Representations of Color Images

Xiaolin Wu and Guangtao Zhai
Department of Electrical and Computer Engineering, McMaster University, Canada
(e-mail: xwu@ece.mcmaster.ca)

***Abstract*—We investigate an intrinsic and useful form of sparsity of color images that was largely overlooked in the literature of image/video processing. This sparsity of multispectral images is revealed and formulated by modeling the image formation process. The underlying new sparse representations of color images are general and can be exploited to improve the performance of existing image restoration algorithms, such as denoising, deblurring, and resolution upconversion.**

Key words: Sparse representations of images, image formation model, image restoration, inverse problem.

I. Introduction

Most digital images and videos are multispectral, having at least three color bands (typically red, green and blue). In natural scenes, most light sources have a continuous spectrum, most objects have a uniform surface material of a certain reflectance and the surface curvature is quite small. As such, different spectral bands of the image signal have high correlations. This research is concerned with sparse representations of multispectral images. The subject is of significance and utility because a wide range of tasks in image processing and computer vision are performed with the assumption, either explicitly or implicitly, that the underlying image signal is sparse.

Recent years have seen a great deal of renewed interests, enthusiasm and progress in sparsity-based image processing, particularly in image restoration. However, quite surprisingly, most published algorithms for image processing and analysis based themselves on the sparsity of luminance component of the image signal and overlooked the sparsities induced by spectral correlations. This leaves a slack in the performance of these algorithms. Mairal *et. al* extended the K-SVD algorithm [1] to color images in the searching of a dictionary based sparse representation of color images [2]. In this paper, to pick up the performance slack we investigate ways to formulate spectral correlations into inherent and computationally amenable sparse representations of multispectral images. Our investigation begins with an image formation model of digital color cameras. This image model and mild assumptions on illumination conditions and imaged objects reveal intrinsic sparsity properties of natural images. It turns out that these sparsities have simple, linear representations that are weighted sum of different spectral bands. This discovery allows the newly revealed sparsities of color images to be readily exploited by an ℓ_1 minimization process, or by linear programming algorithmically. Upon the conclusion of our technical development, it will become self evident how the new results of this paper can be integrated into the general framework of image restoration and used as strong domain knowledge to improve the solution of the corresponding inverse problem.

The remainder of the paper has the following flow of presentation. The image formation model is reviewed in Section II, which leads to the sparse representation that is detailed in Section III. Typical applications of color image denoising and deconvolution are investigated in Section IV. And finally, Section V concludes the paper.

II. Image Formation Model

A multispectral camera records the light reflections of a real world scene. We model the light reflection from the surface of an object by a spectral reflectance function. The surface is associated with a non-negative and bounded reflectance function $f_{u,v}(\lambda)$, where (u, v) denotes the point on the surface that is projected to pixel (u, v), and λ is the wavelength. When this surface point is illuminated by a light source with spectral distribution $L(\lambda)$, the spectral distribution of the reflected light, as observed by the camera, is given by $L(\lambda)f_{u,v}(\lambda)$.

A digital multispectral camera or scanner is equipped with sensors of K different types, $K \geq 3$, each measuring a different spectral sub-band. In consumer electronics, for instance, three spectral subbands, red, green and blue, are commonly used. Type k sensor has its spectral response function $\gamma_k(\lambda)$, $1 \leq k \leq K$. Therefore, the sample value in spectral subband k at pixel position (u, v) is given by

$$x_k(u, v) = \int_\Lambda \gamma_k(\lambda)L(\lambda)f_{u,v}(\lambda)d\lambda \tag{1}$$

where Λ is the spectral range of the camera.

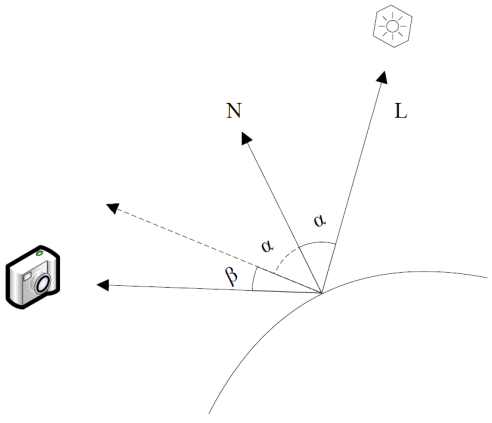

Fig. 1. Illustration of image formation.

The reflectance function $f_{u,v}(\lambda)$ is determined by two factors: the material and the geometry of the object surface, and it is empirically modelled as

$$f_{u,v}(\lambda) = g_{u,v}(\alpha, \beta, \eta) r_{u,v}(\lambda). \quad (2)$$

The term $r(\lambda) \in [0, 1]$ is the reflectance coefficient of the surface material for wavelet length λ, which is independent of surface geometry properties (e.g., curvature and smoothness). The variations in the 2D image of an object in spectral subband k are mostly caused by interactions between the light source $L(\lambda)$ and surface geometry. These interactions can be satisfactorially approximated by

$$g_{u,v}(\alpha, \beta, \eta) = \cos\alpha + \eta \cos^n \beta. \quad (3)$$

The first term $\cos\alpha$ accounts for Lambertian diffuse reflection [3], with α being the angle between the incident light L and the surface normal; the second term $\eta \cos^n \beta$ approximates specular reflection [4], where β is the angle of the reflection ray recorded by the camera at pixel positive (u, v) with the mirror direction of the incident angle (see Fig. 1), and the parameter $\eta \in [0, 1)$ is specular coefficient and $n > 1$ is the degree of surface shininess.

III. Sparse Representations of Spec-Correlation

It follows from (1), (2) and (3) that for each subband

$$x_k(u, v) = (\cos\alpha + \eta \cos^n \beta) \int_\Lambda \gamma_k(\lambda) L(\lambda) r_{u,v}(\lambda) d\lambda. \quad (4)$$

For an object surface S consisting of a uniform material of a fixed $r_{u,v}$, the term $\int_\Lambda \gamma_k(\lambda) L(\lambda) r_{u,v}(\lambda) d\lambda$ remains a constant for all pixels $(u, v) \in S$ on the surface. The 2D waveform of the subband image $\mathbf{x}_k$ corresponding to S is determined by the leading term $(\cos\alpha + \eta \cos^n \beta)$.

First, consider non-specular surfaces (i.e., zero or very small η), which is by far the most common case. Then the dominant term in $(\cos\alpha + \eta \cos^n \beta)$ is $\cos\alpha$. For the light coming from far away (e.g., sun light) and a flat surface S, the angle α is constant everywhere on S. And for a nearby point light source or a surface of small curvature, the angle α changes slowly and smoothly on the surface. Since $\cos\alpha$ can be well approximated by a linear function if α varies in a relatively small range, the subband image $\mathbf{x}_k$ can therefore be modeled well by 2D piecewise linear function in the spatial domain. Next we examine specular surfaces that exhibit highlights, for which $\eta \approx 1$ and $n \gg 1$. Now the dominant term in $(\cos\alpha + \eta \cos^n \beta)$ becomes $\eta \cos^n \beta$. We note that $\cos^n \beta$ can be approximated by two segments of linear function that join at $\beta = 0$. Therefore, the spectral band image $\mathbf{x}_k$ can still be approximately by piecewise linear functions in areas of specular reflections. Summarizing above discussions we conclude that any subband image can be modeled by 2D piecewise linear functions in general, or even by piecewise constant functions on flat surfaces illuminated by distant lights. Consequently, the Laplacian function $\nabla^2 \mathbf{x}_k$, $1 \le k \le K$, offers a natural sparse representation of the spectral band image $\mathbf{x}_k$.

For each spectral band image, at object boundaries the function $\nabla^2 \mathbf{x}_k$ still takes on large nonzero values due to the change of surface materials and/or discontinuities in surface geometry. But we can largely eliminate these large values and create a much sparser representation than each individual Laplacian function $\nabla^2 \mathbf{x}_k$, $1 \le k \le K$, by exploring the spectral correlations. Indeed, it follows from (4) that the ratio between any two spectral bands $\mathbf{x}_j$ and $\mathbf{x}_k$, $\mathbf{x}_k \neq 0$

$$\frac{x_j(u, v)}{x_k(u, v)} = \frac{\int_\Lambda \gamma_j(\lambda) L(\lambda) r_{u,v}(\lambda) d\lambda}{\int_\Lambda \gamma_k(\lambda) L(\lambda) r_{u,v}(\lambda) d\lambda} \quad (5)$$

is constant over pixels $(u, v) \in S$. By factoring out the effects of surface geometry, the ratio image $\mathbf{z}_{j,k} = \mathbf{x}_j / \mathbf{x}_k$ formed by pixelwise division can be approximated by a 2D piecewise constant function. Consequently, the first derivatives $\nabla \mathbf{z}_{j,k}$ of these ratio images, $j \neq k$, $1 \le j, k \le K$, yield $O(K^2)$ sparse representations of the multispectral image $\mathbf{x}$. However, one needs to excise caution if attempted to use the piecewise constant model of $\nabla \mathbf{z}_{j,k}$ in restoration or estimation of $\mathbf{x}$. First, $\nabla \mathbf{z}_{j,k}$ can be numerically unstable due to the risk of division by zero or very small values; second, the nonlinearity of function $\mathbf{z}_{j,k}$ makes the ℓ_1 minimization of $\nabla \mathbf{z}_{j,k}$ computationally very expensive, if not impossible.

To overcome the above difficulties we seek for a linear form of sparse representation of $\mathbf{x}$. By the image formation model of (1) and subsequent discussions, we have

$$\sum_{1 \le k \le K} a_k \mathbf{x}_k(u, v) = (\cos\alpha + \eta \cos^n \beta) F(u, v) \quad (6)$$

where a_k's are the weights and

$$F(u, v) = \int_\Lambda [\sum_{1 \le k \le K} a_k \gamma_k(\lambda)] L(\lambda) r_{u,v}(\lambda) d\lambda. \quad (7)$$

It is easy to see that $F(u, v)$ is a constant over all pixels $(u, v) \in S$, as long as surface S consists of a uniform material. Therefore, any linear combination of $\mathbf{x}_k$, $1 \le k \le K$ can be fit to a piecewise linear function, whose Laplacian yields a sparse representation of the multispectral image. But we will strive for an even stronger sparsity next. Without loss of generality, for $\mathbf{x}_1 \neq 0$, defining

$$b_k = \frac{x_k(u, v)}{x_1(u, v)} = \frac{\int_\Lambda \gamma_k(\lambda) L(\lambda) r_{u,v}(\lambda) d\lambda}{\int_\Lambda \gamma_1(\lambda) L(\lambda) r_{u,v}(\lambda) d\lambda}, 1 \le k \le K, \quad (8)$$

which is a constant for all pixels $(u, v) \in S$, we have

$$\sum_{k=1}^{K} a_k x_k(u, v) = \sum_{k=1}^{K} a_k b_k x_1(u, v). \quad (9)$$

For a pixel location (u, v), $x_1(u, v)$ is a fixed value. By letting $a_k = 1/b_k$, $1 < k \le K$, and substituting it into (9) we have $\sum_{k=1}^{K} a_k x_k(u, v) = K x_1(u, v)$. And by setting $a_1 = 1 - K$, we further have $\sum_{k=1}^{K} a_k x_k(u, v) = 0$. In other words, such weights exist under which the corresponding linear combination of K spectral bands is an all-zero 2D signal within surface S. If we can compute b_k for each surface S, then the linearly combined signal $\mathbf{s} = \sum_{k=1}^{K} a_k \mathbf{x}_k$ with locally

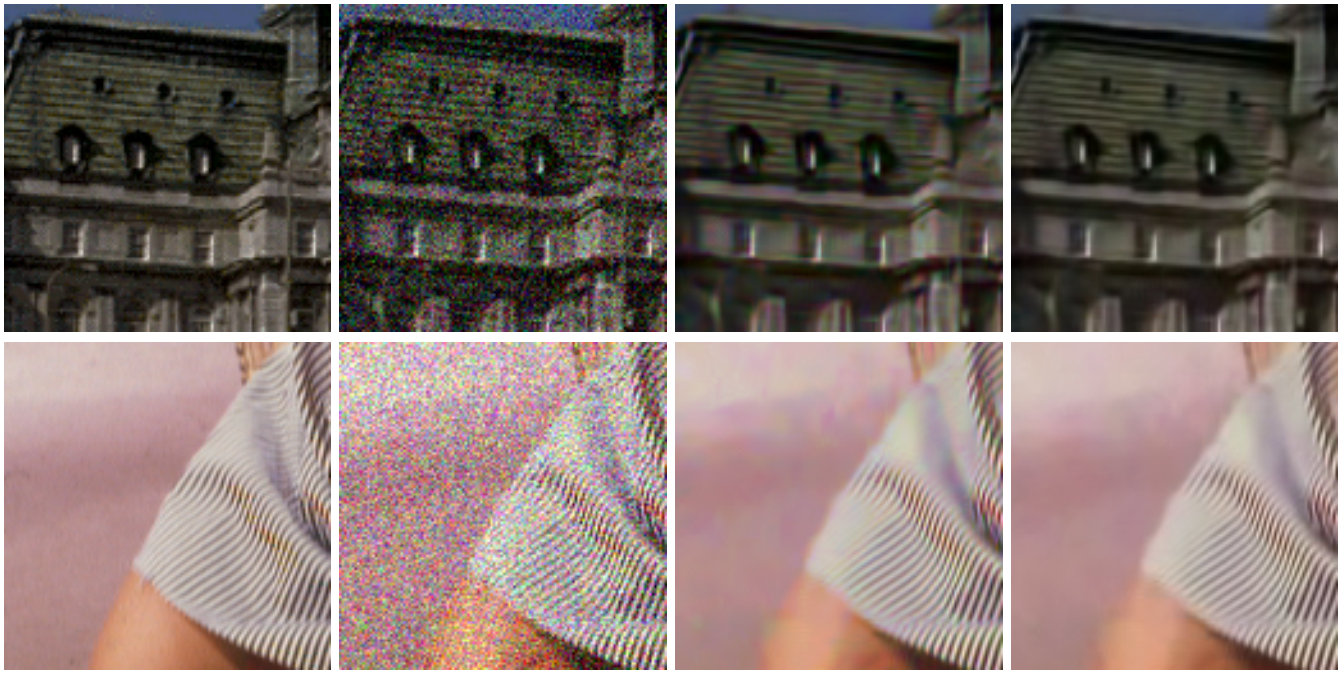

Fig. 2. Example of using the proposed method in image denoising. First row, from left to right: original 'Montreal'; Noisy, $\sigma = 30$, PSNR=19.33 dB; BM3D denoised, PSNR=26.38 dB; Restoration for B3DM with proposed algorithm, PSNR=27.06 dB. Second row, from left to right: original 'Barbara'; Noisy, $\sigma = 35$, PSNR=17.44 dB; BM3D denoised, PSNR=29.06 dB; Restoration for B3DM with proposed algorithm, PSNR=29.97 dB

adaptive weights: $a_1 = 1 - K$, $a_k = 1/b_k$, $1 < k \leq K$, can be made all zero even at the object boundaries. This construction offers the sparsest representation of multispectral images known thus far. Operationally, however, $\gamma(\lambda)$, $L(\lambda)$ and $r(\lambda)$ are generally unknown, direct computation of the ratios b_k with (8) is difficult. But we can estimate the desired weights by setting $a_1 = 1-K$ and solving the following linear least square problem

$$\min_{a_2,a_3,\cdots,a_K} \| \sum_{(u,v)\in S} \sum_{k=1}^{K} a_k x_k(u,v) \|_2. \tag{10}$$

IV. Applications and Experimental Results

In this section, we apply the above introduced linear sparse representations of multispectral images to perform image restoration tasks. In our experiments color images of red, green, and blue spectral bands ($k = R, G, B$) are used. We first compute the linear weights $\hat{a}_R, \hat{a}_G, \hat{a}_B$ for each pixel (u, v) by solving (10) using the least square method. Then the linear combination $\mathbf{x} = \hat{a}_R\mathbf{x}_R + \hat{a}_G\mathbf{x}_G + \hat{a}_B\mathbf{x}_B$ should be a piecewise constant function and therefore its derivative $\nabla\mathbf{x}$ is a sparse representation of the the color image signal. Consequently, we can pose color image restoration as the following constrained ℓ_1 minimization problem:

$$\begin{aligned} &\hat{\mathbf{x}}_R, \hat{\mathbf{x}}_G, \hat{\mathbf{x}}_B = \underset{\mathbf{x}_R,\mathbf{x}_G,\mathbf{x}_B}{\arg\min} \{\|\nabla(\hat{a}_R\hat{\mathbf{x}}_R + \hat{a}_G\hat{\mathbf{x}}_G + \hat{a}_B\hat{\mathbf{x}}_B)\|_{\ell_1}\} \\ &\text{s.t. } \|\mathbf{D}_k\mathbf{x}_k - \mathbf{y}_k\| \leq \varrho_k, |\mathbf{x}_k - \hat{\mathbf{x}}_k| \leq \sigma_k, k \in \{R, G, B\} \end{aligned} \tag{11}$$

In (11), matrix $\mathbf{D}_k$ is the degradation operator for spectral subband k, $k \in \{R, G, B\}$, and vector $\mathbf{y}_k$ is the degraded, observable version of $\mathbf{x}_k$. To further improve the restoration performance, one can add intraband estimates $\hat{\mathbf{x}}_k$'s, if available, as additional constraints. The terms ϱ_k and σ_k are the variance of the noise in spectral subband k and the variance of estimation error of $\hat{\mathbf{x}}_k$, respectively.

We demonstrate the efficacy of the new sparse representations of multispectral images for two of the most common image restoration tasks: denoising and deconvolution. For denoising we use the results of BM3D [5] as $\hat{\mathbf{x}}_k$'s in (11). For deconvolution, $\hat{\mathbf{x}}_k$'s are the results of the Weiner filter. In the our experiments σ denotes the standard deviation of the added white noise while δ is the standard deviation of the Gaussian PSF used in blurring the original image.

Figs. 2 displays some denoising results of BM3D and the proposed sparsity-based technique. In these examples, the latter outperforms the former by up to 0.9 dB. In terms of visual quality, the new technique is noticeably better. In particular, BM3D produces visible color distortions, while our technique is largely free of color artifacts.

As to the application of deconvolution, we show the results of the proposed technique in comparison with those of Weiner filtering in Figs. 3. In this case our technique can outperform the Weiner filter by as much as 1.47 dB. As shown in the figures, the Weiner is prone to speckle color noises, whereas our technique is not.

PSNR results of the denoising and deconvolution applications on more test images are listed in Table I and Table II, where constant improvements over BM3D and Weiner filtering help to further justify this type of new intrinsic sparsity inherited from image-forming process.

Fig. 3. Example of using the proposed method in image deconvolution. First row, from left to right: original 'Oldmill'; Blurry, $\delta = 1$, PSNR=24.78 dB; Wiener filtered, PSNR=26.90 dB; Restoration for Wiener filter with proposed algorithm, PSNR=27.78 dB. Second row, from left to right: original 'Montreal'; Blurry, $\delta = 1$, PSNR=26.08 dB; Wiener filtered, PSNR=27.12 dB; Restoration for Wiener filter with proposed algorithm, PSNR=28.59 dB.

TABLE I
PSNR RESULTS FOR THE APPLICATION OF IMAGE DENOISING OF THE PROPOSED METHOD.

Images	$\sigma = 30$			$\sigma = 35$			$\sigma = 40$		
	Noisy	BM3D	Proposed	Noisy	BM3D	Proposed	Noisy	BM3D	Proposed
oldmill	19.2771	25.5604	**26.5690**	18.0592	24.8357	**25.8734**	17.0208	24.1080	**25.1808**
montreal	19.3385	26.3870	**27.0634**	18.1464	25.7265	**26.4316**	17.1343	25.1278	**25.8846**
barbara	18.7146	29.9650	**30.8075**	17.4457	29.0675	**29.9796**	16.3729	27.8073	**28.8455**
malight	19.4918	29.0211	**29.4611**	18.3046	28.3865	**28.9347**	17.2971	27.7783	**28.4662**
lgthouse	18.7535	28.6445	**29.4496**	17.4983	28.0004	**28.8589**	16.4368	27.3977	**28.3376**
bluheron	18.7979	28.7249	**29.3172**	17.5534	28.1141	**28.7666**	16.4999	27.5227	**28.2817**

TABLE II
PSNR RESULTS FOR THE APPLICATION OF IMAGE DECONVOLUTION OF THE PROPOSED METHOD.

Images	$\delta = 1$			$\delta = 1.5$			$\delta = 2$		
	Blurry	Wiener	Proposed	Blurry	Wiener	Proposed	Blurry	Wiener	Proposed
oldmill	24.7812	26.9056	**27.7863**	22.4223	24.4287	**24.5976**	21.1402	22.6448	**22.6898**
montreal	26.0818	27.1248	**28.5906**	24.0299	25.5091	**25.8356**	22.9178	24.4251	**24.5413**
barbara	23.8686	24.1367	**24.2762**	22.7333	22.9080	**22.9139**	22.4870	22.8384	**22.8582**
malight	27.7038	28.8218	**30.0731**	25.4045	27.5750	**27.7854**	24.0581	25.9088	**25.9294**
lgthouse	28.7066	29.7772	**31.3057**	26.5416	28.6832	**29.0584**	25.1658	27.2509	**27.3446**
bluheron	29.3096	29.1313	**30.7153**	27.5564	29.3128	**29.7061**	26.3600	28.3868	**28.4257**

V. CONCLUSION

A new form of intrinsic sparsity for color images is investigated in this work. This sparsity is deduced from the formation process of multispectral images and study shows that it can be represented as weighted sum of differences between spectral bands. This type of sparsity can be integrated into a general framework of color image restoration. As examples, we describe two typical applications in color image denoising and deconvolution. Substantial visual and PSNR improvements over benchmark methods verify the efficacy of this new sparse representation for color images.

REFERENCES

[1] M. Aharon and M. Elad, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image Processing*, vol. 15, no. 12, p. 3736C3745, 2006.

[2] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Transactions on Image Processing*, vol. 17, no. 1, pp. 53 –69, 2008.

[3] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, pp. 218–233, 2003.

[4] B. T. Phong, "Illumination for computer generated pictures," *Communications of ACM*, vol. 18, no. 6, pp. 311–317, 1975.

[5] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3D transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.