

# Sparse Recovery of Hyperspectral Signal from Natural RGB Images

Boaz Arad and Ohad Ben-Shahar<sup>(✉)</sup>

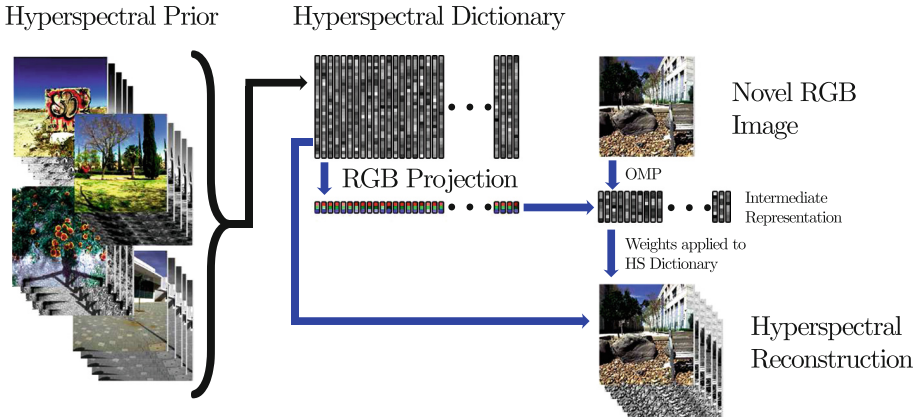
Department of Computer Science, Ben-Gurion University of the Negev,  
Beersheba, Israel

{boazar, ben-shahar}@cs.bgu.ac.il

**Abstract.** Hyperspectral imaging is an important visual modality with growing interest and range of applications. The latter, however, is hindered by the fact that existing devices are limited in either spatial, spectral, and/or temporal resolution, while yet being both complicated and expensive. We present a low cost and fast method to recover high quality hyperspectral images directly from RGB. Our approach first leverages hyperspectral prior in order to create a sparse dictionary of hyperspectral signatures and their corresponding RGB projections. Describing novel RGB images via the latter then facilitates reconstruction of the hyperspectral image via the former. A novel, larger-than-ever database of hyperspectral images serves as a hyperspectral prior. This database further allows for evaluation of our methodology at an unprecedented scale, and is provided for the benefit of the research community. Our approach is fast, accurate, and provides high resolution hyperspectral cubes despite using RGB-only input.

## 1 Introduction

Hyperspectral imagery has been an active area of research since modern acquisition technology became available in the late 1970s [1]. Unlike RGB or multispectral acquisition devices, the goal of hyperspectral imaging is the acquisition of the complete spectral signature reflected from each observable point. The richness of this information facilitates numerous applications, but it also comes with a price – a significant decrease in spatial or temporal resolution (Note that in this sense a typical RGB or other multispectral cameras compromise the third dimension of hyperspectral data, namely the *spectral* resolution.). As a result, the use of Hyperspectral Imaging Systems (HISs) has been limited to those domains and applications in which these aspects of the signal (either spatial, but mostly temporal resolution) were not central – remote sensing (cf. [2]), agriculture (cf. [3]), geology (cf. [4]), astronomy (cf. [5]), earth sciences (cf. [6]), and others. Even in these cases the HIS is often used for the *preliminary* analysis of observable signals in order to characterize the parts of the spectrum that carries valuable information for the application. This information is then used to design multispectral devices (cameras with few spectral bands) that are optimized for that application.



**Fig. 1.** The estimation process: a rich hyperspectral prior is collected, a corresponding hyperspectral dictionary is produced and projected to RGB. Once produced, the dictionary may be used to reconstruct novel images without additional hyperspectral input.

Unlike their use in niche or dedicated applications such as the above, the use of HISs in general computer vision, and in particular in the analysis of natural images, is still in its infancy. The main obstacles are not only the reduced resolution in one of the acquisition “axes” (i.e. spatial, temporal, or spectral), but also the cost of the hyperspectral devices. Both problems result from the attempt to record three dimensional data  $I(x, y, \lambda)$  using two dimensional sensors, which typically require elaborate setups involving some sort of scanning (either spectral or spatial). Ideally, one should obtain a hyperspectral image at high resolution both spatially and spectrally, and do so both quickly (as dictated by the frame rate requirement of the application) and at low cost. While various approximations have been proposed in recent years (see Sect. 2), most require hybrid (and costly) hardware involving both RGB and low resolution hyperspectral measurements. In contrast, here we present a low cost and fast approach requiring only an RGB camera. To address the severely underconstrained nature of the problem (recovering hyperspectral signatures from RGB measurements) we exploit hyperspectral *prior* which is collected and pre-processed only once using tools from the sparse representation literature. As we show, despite the inferior measurements (RGB only vs. RGB endowed with low resolution spectral data), our approach is able to estimate a high quality hyperspectral image, thereby making a significant step toward truly low cost real-time HISs and numerous new scientific and commercial applications.

## 2 Related Work

Acquisition of full spectral signatures has evolved greatly in the last several decades. Originating with spectrometers, nowadays these devices can measure

the intensity of light across a wide range of wavelengths and spectral resolutions (up to picometres) but they lack any form of spatial resolution. Early HISs such as NASA’s AVIRIS [7] produced images with high spatial/spectral resolution using “whisk broom” scanning where mirrors and fiber optics are used to collect incoming electromagnetic signals into a bank of spectrometers pixel by pixel. Newer systems employ “push broom” scanning [8] and utilize dispersive optical elements and light sensitive (e.g., CCD) sensors in order to acquire images line by line. Other systems, often used in microscopy or other lab applications, employ full 2D acquisition through interchangeable filters thus obviating the need for relative motion between the camera and scene at the expense of temporal resolution and high sensitivity to corruption by scene motion. Since purely physical solutions have yet to produce a method for fast acquisition with high spatial and spectral resolution, various methods have been proposed to augment hyperspectral acquisition *computationally*.

Computed tomography imaging spectrometers (CTIS) [9–11] utilize a special diffraction grating to ‘project’ the 3D hyperspectral data cube onto different areas of the 2D imaging sensor. The multiplexed two dimensional data can later be used to reconstruct the hyperspectral cube computationally, but the method as a whole requires both specialized acquisition equipment and significant post processing. Moreover, spatial and spectral resolution is severely limited in relation to sensor size. Building upon advances in the field of compressed sensing, coded aperture HISs [12, 13] and other compressive HS imaging techniques [14] improve upon CTIS in terms of sensor utilization, but still require complex acquisition equipment as well as significant post processing to recover full spectral signatures.

Systems capable of real time acquisition without incurring heavy computational costs have been proposed as well. For example, “Hyperspectral fovea” systems [15, 16] can acquire high resolution RGB data, endowed with hyperspectral data over a small central region of the scene. These systems are mostly useful for applications that require occasional hyperspectral sampling of specific areas rather than a full hyperspectral cube. Du *et al.* [17] proposed a simple prism based system for the acquisition of multispectral video. Unfortunately, this system mandates a direct trade-off between spatial and spectral resolution.

Seeking to improve the spectral and spatial resolution of images acquired from HISs that sample the hyperspectral cube sparsely, Kawakami *et al.* [18] suggested a matrix factorization method in order to obtain high resolution hyperspectral data from input that constitutes both a low resolution hyperspectral image and a high resolution RGB image. Although this method provides high estimation accuracy, it is also extremely computationally intensive, with computational time per image reported in the hours. Assuming the same type of input (high resolution RGB + low resolution spectral image) but replacing some of the extensive matrix factorization computations with simpler propagation methods, Cao *et al.* [19] proposed a specialized hybrid acquisition system capable of producing hyperspectral video at several frames per second.

In more recent studies researchers have increasingly attempted estimation of hyperspectral information using only RGB cameras. By illuminating a target scene with several narrow-band light sources, a process known as “time-multiplexed illumination”, scene reflectance can be estimated across a number of wavelengths. Goel *et al.* [20] proposed such a system capable of estimating 17 spectral bands at 9fps using time multiplexed illumination, while Parmar *et al.* [21] demonstrated the recovery of 31 spectral bands using 5 narrow-band LED sources. This approach seemingly removes computational and temporal hurdles faced by previous efforts but introduces a new constraint of controlled lighting, thus rendering itself ineffective in outdoor conditions, large scale environments or conditions where illumination changes are prohibited.

While single-shot hyperspectral acquisition and hyperspectral video seems within reach, existing systems still require special acquisition hardware and/or complex and costly computations for each frame acquired. The approach we present in this paper improves upon previous work in that the acquisition system that results from it is fast, requires only RGB but no hyperspectral input (and therefore no hyperspectral equipment) whatsoever, and has the bulk of the necessary computations done only once prior to acquisitions.

### 3 Hyperspectral Prior for Natural Images

Key in our work is the exploitation of prior on the distribution of hyperspectral signatures in natural images. In practical terms this prior must be sampled from the real world by acquiring a range of hyperspectral images using a genuine HIS, but this process should be done only once. Naturally, one can use existing collections of hyperspectral images for this purpose. Indeed, databases of reflectance color spectra [22] and images collected from airborne platforms are abundant and readily available for research (NASA’s AVIRIS collection [23] alone contains thousands of images and continues to grow daily). Unfortunately, the former are typically small or limited to specific types of materials while the latter are ill-suited as a prior for ground-level natural images. In the same spirit, however, a collection of *ground-level* hyperspectral images could serve as a prior. To our knowledge only a handful of such data sets have been published to date, with notable examples including those by Brelstaff *et al.* [24] in 1995 (29 images of rural scenes/plant life), by Foster *et al.* [25] in 2002 and 2004 (16 urban/rural scenes), by Yasuma *et al.* [26] (32 studio images of various objects), and by Chakrabarti and Zickler [27] (50 mostly urban outdoor scenes and 27 indoor scenes).

Since collecting hyperspectral image datasets is laborious, most of the above databases are limited in scope (if nothing else, then by the mere number of scenes imaged). At the same time, some of the available data also lacks spatial resolution (for example, the images in the Brelstaff data set are  $256 \times 256$  pixels in size) and all have spectral resolution of 33 channels or less. To allow better collection of hyperspectral prior, and to provide better tools to advance natural hyperspectral imagery research in general, here we provide new and larger



**Fig. 2.** The experimental setup used for the acquisition of our database includes a Specim hyperspectral camera, a computer control rotary stage mounted on a heavy duty tripod, and acquisition computer.

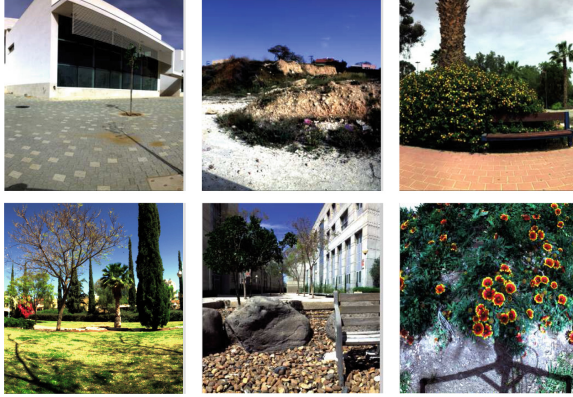
hyperspectral database of natural images captured at high spatial and spectral resolution [28].

Our database of hyperspectral natural images is acquired using a Specim PS Kappa DX4 hyperspectral camera and a rotary stage for spatial scanning (Fig. 2). At this time 100 images were captured from a variety of urban (residential/commercial), suburban, rural, indoor and plant-life scenes (see selected RGB depictions in Fig. 3) but the database is designed to grow progressively. All images are  $1392 \times 1300$  in spatial resolution and 519 spectral bands (400–1,000 nm at roughly 1.25 nm increments). For comparison purposes, and whenever possible, we also compared results using previously published datasets and benchmarks.

## 4 Hyperspectral from RGB

The goal of our research is the reconstruction of the hyperspectral data from natural images from their (single) RGB image. Prima facie, this appears a futile task. Spectral signatures, even in compact subsets of the spectrum, are very high (and in the theoretical continuum, infinite) dimensional objects while RGB signals are three dimensional. The back-projection from RGB to hyperspectral is thus severely underconstrained and reversal of the many-to-one mapping performed by the eye or the RGB camera is rather unlikely. This problem is perhaps expressed best by what is known as *metamerism* [29] – the phenomenon of lights that elicit the same response from the sensory system but having different power distributions over the sensed spectral segment.

Given this, can one hope to obtain good approximations of hyperspectral signals from RGB data only? We argue that under certain conditions this otherwise ill-posed transformation is indeed possible; First, it is needed that the set of hyperspectral signals that the sensory system can ever encounter is confined to a relatively low dimensional manifold within the high or even infinite-dimensional



**Fig. 3.** RGB depictions of few samples from our acquired hyperspectral database.

space of all hyperspectral signals. Second, it is required that the frequency of metamers within this low dimensional manifold is relatively low. If both conditions hold, the response of the RGB sensor may in fact reveal much more on the spectral signature than first appears and the mapping from the latter to the former may be achievable.

Interestingly enough, the relative frequency of metameric pairs in *natural* scenes has been found to be as low as  $10^{-6}$  to  $10^{-4}$  [25]. This very low rate suggests that at least in this domain spectra that are different enough produce distinct sensor responses with high probability. Additionally, repeated findings have been reported to suggest that the effective dimension of visual spectrum luminance is indeed relatively low. Several early studies [30–32] attempted to accurately represent data sets of empirically measured reflectance spectra with a small amount of principal components. While results vary, most agree that 3–8 components suffice to reliably reconstruct the spectral luminance of measured samples. Similar exploration by Hardeberg [33] on several datasets of different pigments and color samples concluded an effective dimension that varies between 13 to 23. Most recently, a similar PCA analysis, though this time on  $8 \times 8$  tiles from the Chakrabarti dataset, found that the first 20 principle components account for 99 % of the sample variance [27]. This last result is of additional interest since it implies that hyperspectral data in the visual spectrum is sparse both spectrally and spatially.

One may argue that the sparsity of natural hyperspectral signatures is to be expected. Indeed, the spectral reflectance of an object is determined by two main factors: its material composition and the spectral properties of the illumination. While many factors may affect the spectrum reflected by a material sample in subtle ways, it can be generally viewed as a *linear* combination of the reflected spectra produced by the different materials composing the sample [34]. Although the range of possible materials in nature may be large, it is conceivable to assume that only few contribute to the spectrum measured *at each*

*particular pixel* in the hyperspectral image. Hence, a natural way to represent spectra observed in natural images is a sparse combination of basis spectra stored in a dictionary. Indeed, among several methods proposed in the field of color science for reflectance estimation from RGB images [35], *regression estimation* suggests the use of a dictionary containing a collection of reflectance/measurement pairs in order to estimate the underlying reflectance of new measurements. While previous studies [21, 36, 37] have attempted to apply the regression estimation method for reflection estimation, most of them were limited to theoretical studies on small datasets of known “generic” spectra (such as the Munsell color chip set) or to domain specific tasks [36]. Despite their limited scope, these studies indicate that accurate spectral recovery may be achieved from RGB data. Further optimism may be garnered from the recent work of Xing *et al.* [38] demonstrating noise reduction and data recovery in hyperspectral images based on a sparse spatio-spectral dictionary. Although based upon aerial imagery, Xing’s results demonstrate the power of sparse representations and over-complete dictionaries in hyperspectral vision.

#### 4.1 Spectra Estimation via Sparse Dictionary Prior

Building upon the observed sparsity of natural hyperspectral images, we suggest a sparse dictionary reconstruction approach based on a rich hyperspectral prior for reconstruction of hyperspectral data from RGB measurements. First, a rich hyperspectral prior is collected, preferably (but not necessarily) from a set of domain specific scenes. This prior is then reduced computationally to an over-complete dictionary of hyperspectral signatures. Let  $D_h$  be such an overcomplete dictionary  $\mathbf{h}_i$  (expressed as column vectors) in natural images:

$$D_h = \{\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n\}. \quad (1)$$

Once obtained, the dictionary is projected to the sensor space via the receptor spectral absorbance functions. While this formulation is general and suits different types of sensors, here we focus on RGB sensors and the RGB response profiles. If  $d = \dim(\mathbf{h}_i)$  is the dimension of the spectral signatures after quantization to the desired resolution, these projections are expressed as inner products with matrix  $R$  of dimensions  $3 \times d$  which yields a corresponding RGB dictionary  $D_{rgb}$

$$D_{rgb} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\} = R \cdot D_h. \quad (2)$$

of three dimensional vectors  $\mathbf{c}_i = (r_i, g_i, b_i)^T$  such that

$$\mathbf{c}_i = R \cdot \mathbf{h}_i \quad \forall \mathbf{c}_i \in D_{rgb}. \quad (3)$$

The correspondence between each RGB vector  $\mathbf{c}_i$  and its hyperspectral originator  $\mathbf{h}_i$  is maintained for the later mapping from RGB to hyperspectral signatures. This also completes the pre-processing stage which is done only once.

Given an RGB image, the following steps are used to estimate the corresponding hyperspectral image of the scene. For each pixel query  $\mathbf{c}_q = (r_q, g_q, b_q)^T$  encountered in the RGB image, a weight vector  $\mathbf{w}$  is found such that:



$$D_{rgb} \cdot \mathbf{w} = \mathbf{c}_q. \quad (4)$$

The weight vector  $\mathbf{w}$  must adhere to the same degree of sparsity imposed on  $D_h$  at the time of its creation. Once  $\mathbf{w}$  is found, the spectrum  $\mathbf{h}_q$  underlying  $\mathbf{c}_q$  is estimated by the same linear combination, this time applied on the hyperspectral dictionary:

$$\mathbf{h}_q = D_h \cdot \mathbf{w}. \quad (5)$$

Since  $D_{rgb}$  was generated from  $D_h$  it follows (from Eqs. 2 and 4) that the reconstructed spectrum is consistent with the dictionary:

$$\mathbf{c}_q = R \cdot \mathbf{h}_q. \quad (6)$$

but whether or not  $\mathbf{h}_q$  is indeed an accurate representation of the hyperspectral data that generated the pixel  $\mathbf{c}_q$  depends on the representational power of the dictionary and must be tested empirically. As is demonstrated in Sect. 5, reconstruction quality is directly affected by the scope and specificity of the hyperspectral prior.

## 5 Implementation and Results

Our hyperspectral recovery method was tested using images from our newly acquired hyperspectral database (cf. Sect. 3). The spectral range used from each image was limited roughly to the visual spectrum and computationally reduced via proper binning of the original narrow bands to 31 bands of roughly 10 nm in the range 400–700 nm. This was done both to reduce computational cost but mostly to facilitate comparisons to previous benchmarks that employ such representation.

To test the proposed algorithm we selected a test image from the database and mapped it to RGB using CIE 1964 color matching functions. 1000 random samples from each of the *remaining* images were then combined to create the over complete hyperspectral dictionary  $D_h$  using the K-SVD algorithm [39]. The dictionary size was limited to 500 atoms, under a sparsity constraint of 28 non-zero weights per atom. These parameters were determined to be ideal via exploration of the parameter space. Figure 6b depicts performance over variable parameters, demonstrating the robustness of our method to parameter selection.

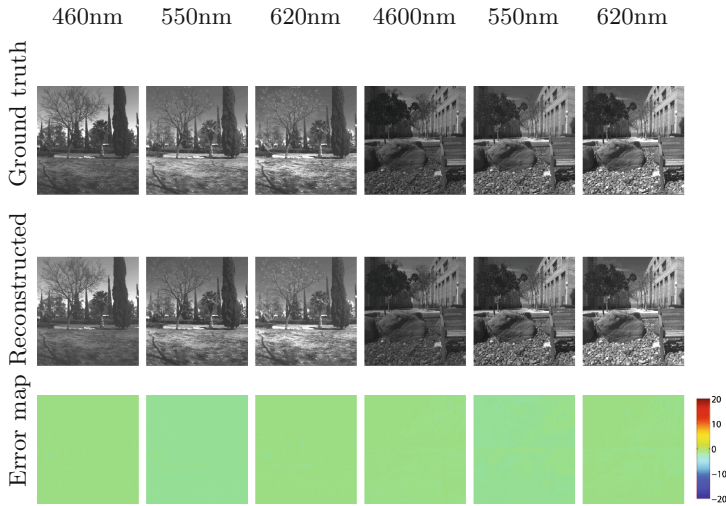
The resulting dictionary was then projected to RGB to form  $D_{rgb}$ . Once all these components have been obtained, the hyperspectral signature of each pixel of the test image was estimated as described above, where the dictionary representation of each RGB pixel was computed with the Orthogonal Match Pursuit (OMP) [40] algorithm.

The process just described was repeated until each image had been selected for testing and independently reconstructed several times to discount the stochastic aspect of the dictionary. The reconstructed hyperspectral images were compared to ground-truth data from the database and RMSE errors were computed. Additionally, we repeated the same process for specific image subsets in the database (urban scenes, rural scenes, etc.) in order to explore the effect of domain-specific prior on reconstruction performance.



## 5.1 Experimental Results

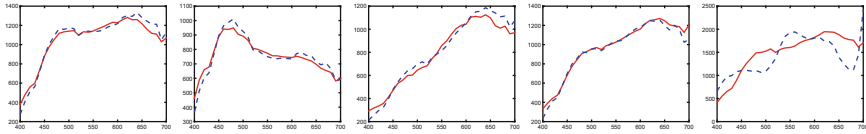
Figure 5 exemplifies the quality of spectra reconstruction obtained with our approach (recall that the only input during reconstruction is the RGB signal). This type of results, that represent not only qualitative but also very accurate quantitative reconstructions, characterizes the vast majority of pixels in all images in the database. Figure 4 shows a comparison of the reconstructed and ground truth spectral bands for two selected images. Notice the relatively shallow error maps (using the same scale as used in Kawakami *et al.* [18] for comparison).



**Fig. 4.** Comparison of reconstructed luminance to ground truth luminance in selected spectral bands of two images (cf. Fig. 3). Luminance error presented on a scale of  $\pm 255$  (as presented in Kawakami *et al.* [18]).

Estimation errors were reported in terms of luminance error divided by ground truth luminance, thus preventing a bias towards low errors in low-luminance pixels. Additionally, absolute RMSE values were reported on a scale of 0–255 in order to facilitate comparison to results reported in previous work. Table 1 presents pooled results from the evaluation process described above while Fig. 6a displays the average RMSE per spectral channel of reconstructed images. On average across our entire database, hyperspectral images were reconstructed with a relative RMSE error of 0.0756. Errors were mostly pronounced in channels near the edge of the visible spectrum. As the table further shows, when both dictionary construction and the reconstruction procedures are restricted to specific domains, performance typically improves even further since images from a certain category are more likely to share hyperspectral prior. It is therefore expected that the suggested methodology will perform especially well in restricted domain tasks. Conversely, cross-domain tests (i.e. reconstruction of

images from the “Park” set using a dictionary generated from the “Rural” set) produced comparable RMSE values to the reconstructions with a general prior, indicative that such dictionaries may be useful across various domains.



**Fig. 5.** A sample failure case (right) and 4 random samples (left) of spectra reconstruction (blue, dashed) vs. ground truth data (red). (Color figure online)

Finally, we applied our approach to the hyperspectral database acquired by Chakrabarti and Zickler [27]. Dividing the set to indoor and outdoor images, average RMSE over each of these subsets is reported at the bottom of Table 1. Compared to results on our database, performance is degraded. The indoor subset exhibited low absolute RMSE values, alongside high relative RMSE values - indicating that reconstruction errors were largely constrained to low-luminance pixels, which are indeed abundant in the subset. Degraded performance is further explained by the fact that the Chakrabarti database was sampled in the 420–720 nm range, outside the 400–700 nm effective range of the CIE color response function. Additionally some hyperspectral blurring was found to contaminate the data. Indeed, while Chakrabarti and Zickler [27] provided motion masks for scene segments suspected with extensive motion, more subtle motions that are not captured by these masks are observable and may affect the results. Note that even in the case of the indoor subset, absolute RMSE values are comparable to previous reported results (e.g. Kawakami *et al.* [18]).

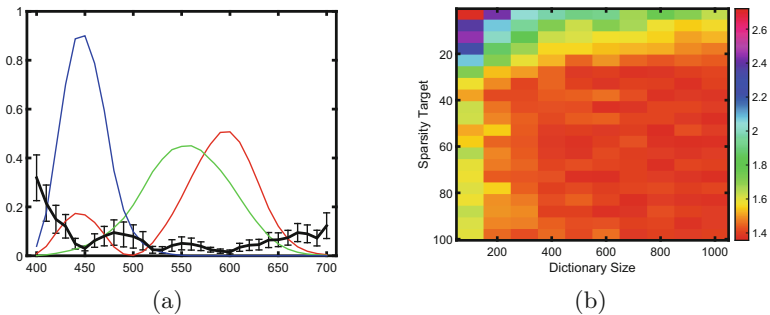
## 5.2 Comparison to Prior Art

Since previous work on hyperspectral evaluation differs either in input (RGB+HS vs. RGB only) or evaluation scale (ranging between 102 pixels in Parmar *et al.* [21] and  $10^6$  pixels in Kawakami *et al.* [18] vs. over  $10^8$  reconstructed pixels presented here) it may be difficult to make an equal-ground comparison. Nevertheless, we have compared our approach to results presented by Kawakami *et al.* [18] and tested our algorithm on the Yasuma data set [26]. Sadly, while the method presented by Parmar *et al.* [21] (cf. Sect. 4) may be applied to three-channel input, their paper only presented two data-points reconstructed from 8-channel input thus rendering comparison impossible.

As noted earlier, the Yasuma data set constitutes 32 studio images, many of which contain large dark background areas. Naïve acquisition of our hyperspectral prior by randomly sampling these images is likely to produce a biased dictionary where the genuine hyperspectral information is severely underrepresented. Additionally, being an indoor collection of different random objects, it is

**Table 1.** Average relative/absolute root mean square error of reconstruction over different image sets. Absolute RMSE values are shown in the range of 8-bit images (0–255).

Data Set	Relative RMSE	Absolute RMSE
<b>Complete data set</b>	<b>0.0756</b>	<b>2.633</b>
Park subset	0.0589	2.105
Indoor subset	0.0507	1.417
Urban subset	0.0617	2.260
Rural subset	0.0354	1.582
Plant-life subset	0.0469	1.816
<b>Cross domain</b>		
Park subset from rural prior	0.0801	2.693
Rural subset from park prior	0.0592	3.121
<b>Chakrabarti data set</b>		
Outdoor subset	0.1231	3.466
Indoor subset	0.5207	5.685

**Fig. 6.** (a) Average relative RMSE per channel across reconstructions (black). CIE 1964 color matching functions, displayed at an arbitrary scale, overlaid for reference (red, green, blue). (b) Reconstruction RMSE over a subset of reconstructed images as a function of parameter selection. Note that performance peaks quickly for sparsity target around 30 and dictionary size around 500 (and remain relatively stable thereafter). (Color figure online)

unlikely that a prior collected from one could be used successfully to reconstruct spectral signatures for others. To overcome these limitations, a hyperspectral prior was sampled from each image separately before reconstruction. 10,000 pixels (3.8% of each image) were sampled either randomly from the entire image or from a central region of the image to avoid the dark (hyperspectrally poor) background (if existed). These were then reduced computationally to a hyperspectral dictionary. Additionally, initial atoms for the K-SVD algorithm were selected either randomly from the sampled prior, or via maximization of the

distance between their projected RGB values. Reconstructions were performed using each of the resulting dictionaries and the results are reported in Table 2.

**Table 2.** Numerical comparison of root mean squared error between methods. The numbers are shown in the range of 8-bit images (0–255) in order to match results presented by Kawakami *et al.* [18]. Note the comparable results of our method despite using much inferior image data (RGB+hyperspectral prior vs. RGB+low resolution hyperspectral of each image) during the reconstruction.

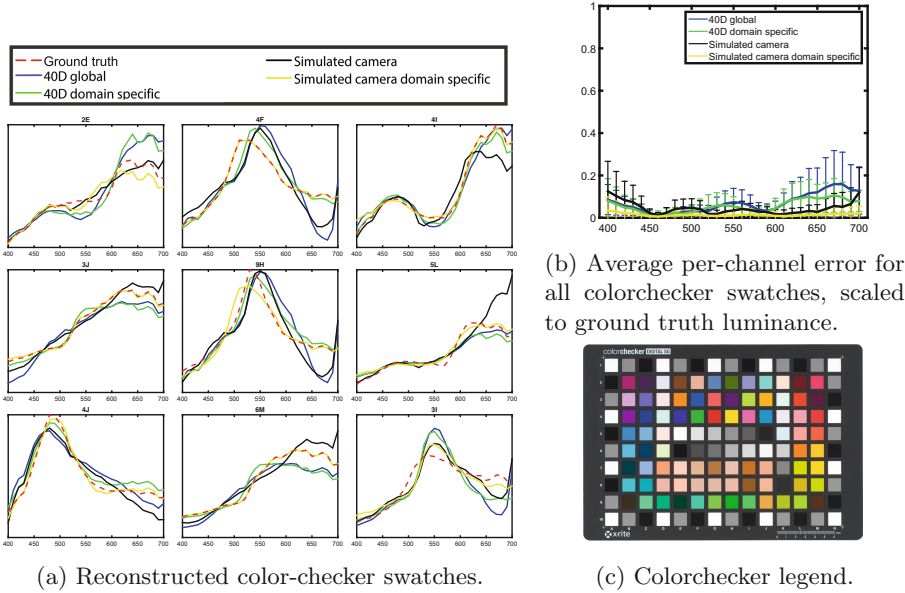
	Ours	Kawakami [18]
Dataset Avg	<b>5.4</b>	not reported
Balloons	5.2	<b>3.0</b>
Beads	<b>5.3</b>	9.2
Sponges	<b>3.1</b>	3.7
Oil Painting	<b>3.3</b>	4.7
Flowers	<b>4.2</b>	5.4
CD	8.7	<b>8.2</b>
Fake/Real Peppers	5.7	<b>4.7</b>
Photo and Face	<b>3.3</b>	<b>3.3</b>

As can be observed in the table, despite using only RGB for reconstruction, results are comparable (note that Kawakami *et al.* [18] reported results only on 8 images out of the entire database). Importantly, while Kawakami *et al.* [18] reported computation of several *hours* for factorization and reconstruction of a  $4008 \times 2672$  image on an eight-core CPU, our algorithm completed both dictionary construction and image reconstruction in seconds (timed on a modest four-core desktop using Matlab implementation). Needless to say that our approach can be massively parallelized in a trivial way since the reconstruction of each pixel is independent of the others. Video rate reconstruction is therefore well within reach.

### 5.3 Reconstruction from Consumer RGB Camera

The eventual goal of our research is the ability to turn consumer grade RGB cameras into a hyperspectral acquisition devices, thus permitting truly low cost and fast HISs.

To demonstrate the feasibility of our methodology, spectra from a color calibration target (X-Rite ColorChecker Digital SG c.f. Fig. 7c) was reconstructed using RAW sensor output recorded from an unmodified consumer camera (Canon 40D). Since a calibrated hyperspectral prior is key in successful reconstruction, the camera filter response profiles must be known. While most manufacturers do not provide this information, Jiang *et al.* [41] have estimated the response



**Fig. 7.** (a) Reconstructed color-checker swatches. (b) Average per-channel error for all colorchecker swatches, scaled to ground truth luminance. (c) Colorchecker legend.

**Table 3.** Average relative root mean square over all colorchecker swatches. Real-world/simulated camera data reconstructed using a domain specific (sampled from HS colorchecker image) and global (sampled from many natural images) dictionary.

Camera	Global prior	Domain specific prior
Canon 40D	0.0757	0.557
Simulated Canon 40D	0.0596	0.0149

profile of several cameras empirically. Using these experimental response functions we created dictionaries with the prior being either the entire database (dubbed “global” in Table 3) or just a hyperspectral image of the calibration target (representing “domain-specific” prior). Spectra were reconstructed from both the real 40D camera and a simulated one (whose response was computed by applying the experimental response functions to the hyperspectral image).

Prior to reconstruction, some disagreement was found between actual camera response and the response predicted by applying the empirical response function to acquired HS information. Average relative RMSE error of empirical camera response vs. expected response was 0.0474. Several factors may contribute to these discrepancies including: chromatic aberrations induced by the camera lens, noise or non-linearity in the camera sensor, and manufacturing variability of the sensor and/or Bayer filters. Selected results are presented in Fig. 7a. Although the reconstruction dictionary was based on an imperfect response function, the

average reconstruction error across all color-checker swatches was comparable to simulated results (c.f. Table 3) with most errors constrained to the far ends of the visible spectrum (c.f. Fig. 7b) where, again, typical RGB filters provide little to no information.

## 6 Implications and Summary

As is evident from the method and results we just introduced, both RGB samples and their corresponding reconstructed spectra, are almost always well represented by 3 dictionary atoms. This may seem expected when it comes to the RGB samples themselves<sup>1</sup>. But why this works so well for the hyperspectral signatures may be a far greater surprise. This largely empirical finding may in fact explain the disagreement between previous works regarding the effective dimensionality of natural image spectra (c.f. Sect. 4), as one may conclude that the dimensionality of this spectral space relies heavily on basis selection. While the stability of RGB-spectra mapping may depend on the low abundance of metamers in both training and test images (and indeed, in nature itself), our experimental results show that it is robust across variable outdoor illumination conditions and scenes. Clearly, the issue of metamers deserves a deeper look that is outside the scope of this paper, and is part of extensive past and future research.

In summary, we have presented a computational approach for the reconstruction of high resolution hyperspectral images from RGB-only signals. Our method is based on collecting hyperspectral prior (either general or domain specific) for the construction of sparse hyperspectral dictionary, whose projection into RGB provides a mapping between RGB atoms to hyperspectral atoms. Describing an arbitrary RGB signal as a combination of RGB atoms then facilitates the reconstruction of the hyperspectral source by applying the same combination on the corresponding hyperspectral atoms. Experimental evaluation, unprecedented in its scope, has demonstrated how our approach provides comparable results to hybrid HS-RGB systems despite relying on significantly inferior data for each image (RGB only vs RGB + low resolution hyperspectral in previous approaches) during the construction phase, thus leading the way for turning consumer grade RGB cameras into full fledged HIS. Towards this end we have also provided a progressively growing large scale database of high resolution (both spatially and spectrally) images for the use of the research community.

**Acknowledgments.** This research was supported in part by the by the Israel Science Foundation (ISF FIRST/BIKURA Grant 281/15) and the European Commission (Horizon 2020 grant SWEEPER GA no. 644313). We also thank the Frankel Fund and the Helmsley Charitable Trust through the ABC Robotics Initiative, both at Ben-Gurion University of the Negev.

---

<sup>1</sup> After all, RGB samples are indeed 3 dimensional and it is quite unlikely, and never happened in our experiments, that  $D_{rgb}$  will not span RGB space  $\mathbb{R}^3$ .

## References

1. Kerekes, J., Schott, J.: Hyperspectral imaging systems. Hyperspectral data exploitation: theory and applications (2007)
2. Lillesand, T., Kiefer, R., Chipman, J., et al.: Remote Sensing and Image Interpretation. Wiley, New York (2004)
3. Haboudane, D., Miller, J., Pattey, E., Zarco-Tejada, P., Strachan, I.: Hyperspectral vegetation indices and novel algorithms for predicting green lai of crop canopies: modeling and validation in the context of precision agriculture. In: Remote Sensing of Environment (2004)
4. Cloutis, E.: Review article hyperspectral geological remote sensing: evaluation of analytical techniques. *Int. J. Remote Sens.* **17**, 2215–2242 (1996)
5. Hege, E., O’Connell, D., Johnson, W., Basty, S., Dereniak, E.: Hyperspectral imaging for astronomy and space surveillance. In: SPIE (2004)
6. Mustard, J., Sunshine, J.: Spectral analysis for earth science: investigations using remote sensing data. In: Manual of Remote Sensing, Remote Sensing for the Earth Sciences (1999)
7. Green, R., Eastwood, M., Sarture, C., Chrien, T., Aronsson, M., Chippendale, B., Faust, J., Pavri, B., Chovit, C., Solis, M., Olah, M., Williams, O.: Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS). In: Remote Sensing of Environment (1998)
8. James, J.: Spectrograph Design Fundamentals. Cambridge University Press, New York (2007)
9. Descour, M., Dereniak, E.: Computed-tomography imaging spectrometer: experimental calibration and reconstruction results. *Appl. Opt.* **34**, 4817–4826 (1995)
10. Okamoto, T., Yamaguchi, I.: Simultaneous acquisition of spectral image information. *Opt. Lett.* **16**, 1277–1279 (1991)
11. Johnson, W., Wilson, D., Bearman, G.: Spatial-spectral modulating snapshot hyperspectral imager. *Appl. Opt.* **45**, 1898–1908 (2006)
12. Brady, D., Gehm, M.: Compressive imaging spectrometers using coded apertures. In: Defense and Security Symposium (2006)
13. Gehm, M., John, R., Brady, D., Willett, R., Schulz, T.: Single-shot compressive spectral imaging with a dual-disperser architecture. *Opt. Express* **15**, 14013–14027 (2007)
14. Lin, X., Wetzstein, G., Liu, Y., Dai, Q.: Dual-coded compressive hyperspectral imaging. *Opt. Lett.* **39**, 2044–2047 (2014)
15. Fletcher-Holmes, D., Harvey, A.: Real-time imaging with a hyperspectral fovea. *J. Opt. A Pure Appl. Opt.* **7**, S298–S302 (2005)
16. Wang, T., Zhu, Z., Rhody, H.: A smart sensor with hyperspectral/range fovea and panoramic peripheral view. In: CVPR (2009)
17. Du, H., Tong, X., Cao, X., Lin, S.: A prism-based system for multispectral video acquisition. In: ICCV (2009)
18. Kawakami, R., Wright, J., Yu-Wing, T., Matsushita, Y., Ben-Ezra, M., Ikeuchi, K.: High-resolution hyperspectral imaging via matrix factorization. In: CVPR (2011)
19. Cao, X., Tong, X., Dai, Q., Lin, S.: High resolution multispectral video capture with a hybrid camera system. In: CVPR (2011)
20. Goel, M., Whitmire, E., Mariakakis, A., Saponas, T.S., Joshi, N., Morris, D., Guenter, B., Gavrilu, M., Borriello, G., Patel, S.N.: Hypercam: hyperspectral imaging for ubiquitous computing applications (2015)



21. Parmar, M., Lansel, S., Wandell, B.A.: Spatio-spectral reconstruction of the multispectral datacube using sparse recovery. In: ICIP (2008)
22. Kohonen, O., Parkkinen, J., Jääskeläinen, T.: Databases for spectral color science. *Color Res. Appl.* **31**, 381–390 (2006)
23. NASA: Airborne Visual Infrared Imaging Spectrometer website. <http://aviris.jpl.nasa.gov/>
24. Brelstaff, G., Párraga, A., Troscianko, T., Carr, D.: Hyperspectral camera system: acquisition and analysis. In: SPIE (1995)
25. Foster, D., Amano, K., Nascimento, S., Foster, M.: Frequency of metamerism in natural scenes. *JOSA A* **23**, 2359–2372 (2006)
26. Yasuma, F., Mitsunaga, T., Iso, D., Nayar, S.: Generalized assorted pixel camera: post-capture control of resolution, dynamic range and spectrum. Technical report (2008)
27. Chakrabarti, A., Zickler, T.: Statistics of real-world hyperspectral images. In: CVPR (2011)
28. BGU Interdisciplinary Computational Vision Laboratory (iCVL): Hyperspectral Image Database website. <http://www.cs.bgu.ac.il/~icvl/hyperspectral/>
29. Palmer, S.: *Vision Science: Photons to Phenomenology*. The MIT Press, Cambridge (1999)
30. Cohen, J.: Dependency of the spectral reflectance curves of the Munsell color chips. *Psychonomic Sci.* **1**, 369–370 (1964)
31. Maloney, L.: Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *JOSA A* **3**, 1673–1683 (1986)
32. Parkkinen, J.P., Hallikainen, J., Jaaskelainen, T.: Characteristic spectra of Munsell colors. *JOSA A* **6**, 318–322 (1989)
33. Hardeberg, J.Y.: On the spectral dimensionality of object colors. In: *Proceedings of CGIV 2002, First European Conference on Colour in Graphics* (2002)
34. Adams, J., Smith, M., Gillespie, A.: Simple models for complex natural surfaces: a strategy for the hyperspectral era of remote sensing. In: *IGARSS* (1989)
35. Heikkinen, V., Lenz, R., Jetsu, T., Parkkinen, J., Hauta-Kasari, M., Jääskeläinen, T.: Evaluation and unification of some methods for estimating reflectance spectra from RGB images. *JOSA A* **25**, 2444–2458 (2008)
36. López-Álvarez, M.A., Hernández-Andrés, J., Romero, J., Olmo, F., Cazorla, A., Alados-Arboledas, L.: Using a trichromatic CCD camera for spectral skylight estimation. *Appl. Opt.* **47**, 31–38 (2008)
37. Ayala, F., Echávarri, J.F., Renet, P., Negueruela, A.I.: Use of three tristimulus values from surface reflectance spectra to calculate the principal components for reconstructing these spectra by using only three eigenvectors. *JOSA A* (2006)
38. Xing, Z., Zhou, M., Castrodad, A., Sapiro, G., Carin, L.: Dictionary learning for noisy and incomplete hyperspectral images. *SIAM J. Imaging Sci.* **5**, 33–56 (2012)
39. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: an algorithm for designing over-complete dictionaries for sparse representation. *IEEE Trans. Signal Process.* **54**, 4311–4322 (2006)
40. Pati, Y., Rezaifar, R., Krishnaprasad, P.: Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. In: *The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers* (1993)
41. Jiang, J., Liu, D., Gu, J., Susstrunk, S.: What is the space of spectral sensitivity functions for digital color cameras? In: *WACV, IEEE* (2013)

Computer Vision – ECCV 2016

14th European Conference, Amsterdam, The

Netherlands, October 11–14, 2016, Proceedings, Part

VII

Leibe, B.; Matas, J.; Sebe, N.; Welling, M. (Eds.)

2016, XXIX, 869 p. 332 illus., Softcover

ISBN: 978-3-319-46477-0