This ICCV paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the version available on IEEE Xplore.

Filter Selection for Hyperspectral Estimation

Boaz Arad, Ohad Ben-Shahar
Department of Computer Science
Ben-Gurion University of the Negev
boazar@cs.bgu.ac.il, ben-shahar@cs.bgu.ac.il

Abstract

While recovery of hyperspectral signals from natural RGB images has been a recent subject of exploration, little to no consideration has been given to the camera response profiles used in the recovery process. In this paper we demonstrate that optimal selection of camera response filters may improve hyperspectral estimation accuracy by over 33%, emphasizing the importance of considering and selecting these response profiles wisely. Additionally, we present an evolutionary optimization methodology for optimal filter set selection from very large filter spaces, an approach that facilitates practical selection from families of customizable filters or filter optimization for multispectral cameras with more than 3 channels.

1. Introduction

In recent years, recovery of Hyperspectral (HS) images from consumer or low cost equipment has become a field of growing interest. Several works have demonstrated the recovery of whole-scene HS information via hybrid HS-RGB systems [16, 15, 5, 8], RGB cameras endowed with controlled illumination [10, 23, 24, 6], multiple RGB cameras [21], and software-only reconstruction from single RGB images [4, 20].

Although high accuracy HS estimation has indeed been demonstrated using only consumer-grade RGB cameras [4, 20, 20], not all RGB cameras are created equal as their spectral response functions may vary significantly from each other [13, 11]. In fact, camera response functions vary not only across manufacturers or camera classes, but often even between generations of similar camera models (i.e. iPhone 6 vs. iPhone 7, Canon 70D vs. Canon 80D). It has previously been shown that the effective dimension of camera spectral response profiles is at least 8 per channel [21], Fig. 1 illustrates this variability for one set of response functions. In this paper, we explore the effects of camera response functions on HS estimation performance and present a methodology for efficient filter selection in RGB/Multi-Spectral systems aimed at HS-reconstruction.

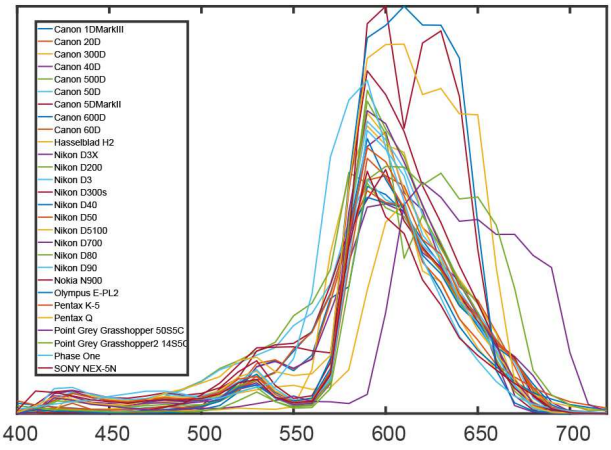


Figure 1: Normalized response function for the red channel of 28 different consumer cameras, as measured by Jiang *et al*. [13]. Note that while most responses are qualitatively similar, there is a high degree of variability among them.

2. Hyperspectral Estimation from RGB

Hyperspectral images depict entire scenes at a high spectral resolution. This additional spectral information provides advantages over the three bands acquired by RGB cameras, but does not come without cost. Traditional HS imaging systems employ either spatial scanning (known as "push-broom" scanners) or temporal scanning (filter wheel systems). "Push-broom" systems are well-suited for airborne acquisition, where scene scanning may be achieved by relative motion between it and the imaging platform. Conversely, filter wheel systems are well suited for laboratory settings where imaging targets can be held in a stationary position for the duration of acquisition.

Unfortunately, neither "push-broom" nor filter wheel systems are well suited for imaging in contexts where most *consumer* cameras are used today. If any in-scene motion occurs during acquisition, the former will produce spa-

tial distortions in the image, while the latter will produce spectral distortions. In order to overcome these limitations, various approaches have been attempted. For example: computed tomography systems [7, 22, 14] offer high accuracy "single shot" HS images, at the cost of significantly reduced resolution relative to sensor size. Alternatively "Hyperscptral fovea" systems [9, 26] produce high-resolution RGB images augmented with hyperspectral information from a small, central area of the scene. Researchers have also begun using low-resolution HS information, coupled with high-resolution RGB information in order to extrapolate high-resolution HS information from the entire scene [16, 15, 5, 8].

More recently, it has been demonstrated that a single RGB image may suffice to produce accurate estimates of HS information over an entire image [4, 20]. Such single shot, RGB/consumer camera-based hyperspectral acquisition systems provide significant advantages over previous approaches. Clearly, rapid acquisition allows imaging of moving targets as well as hyperspectral video at full sensor resolution while the use of consumer/RGB sensors significantly reduces system costs. Furthermore, as they lack the complex optics necessary for hybrid acquisition systems, HS-from-RGB systems allow for easy integration in existing form factors such as cellular-phone cameras, SLR cameras, drones etc. Owing to these advantages, this paper focuses on such single-shot HS-from-RGB estimation methods that do not rely on controlled illumination and/or true HS input. Our goal is to explore the dependency of these HS-from-RGB/Multi-Spectral reconstruction systems on their spectral filters and to suggest methods for filter selection to optimize their performance.

3. Filter Selection

When designing a camera system, several optical filters can be selected in order to shape its response to specific wavelengths. We call this set of optical filters the *filter set* of the camera system . RGB cameras, for example, have filter sets of size 3, whose filters loosely approximate the CIE-1931 color matching functions [19, 13]. Filter sets of size n may be selected from some *filter space* - a finite (or infinite) space describing all possible combinations of n filters from a finite (or infinite) given set of filters.

The choice of a specific filter set from a given filter space clearly affects the camera systems sensor response. But as demonstrated later in Sec. 5, filter selection can profoundly impact the HS estimation performance of a camera system as well. In the case of "off-the-shelf" or consumer equipment, a reasonable strategy for selecting an optimal camera response profile may be exhaustive search. Evaluating the expected performance of hundreds or even thousands of candidate cameras may be laborious but as long as their response functions are known - it is a straightforward process which can be completed in a relatively short time.

Despite the diversity of response functions among consumer camera (cf. Fig 1) there may be much to gain by exploring additional classes of filter sets, including those that differ significantly from the CIE-1931 color matching functions. Filter set optimization the has been studied partially as part of "optimal band selection" in the field of remote sensing but remains an open problem [12, 27, 18, 17]. Moreover, even if theoretically ideal filter sets could be easily computed - they may prove too complex to implement as optical coatings or in a Bayer filter mosaic. Hence in order to enjoy the advantages of an improved filter set within a low-cost system, filters must be selected from the domain of commercially available, or inexpensively manufacturable filters. Unfortunately, this constraint hardly reduces the complexity of the problem, as thousands of such filters are available through various optics vendors (the OMEGA Optical company catalog[2] alone contains over 1000 filters). Additionally, many manufacturers are able to produce customized filters within specific sets of parameters at a relatively low cost [1].

Even while considering only filter sets of size 3 selected from 1000 "off-the-shelf" filters, the size of the filter space quickly balloons to $\binom{1000}{3} = \mathcal{O}(10^8)$, making exhaustive search unreasonable. The number of possible combinations increases by several more orders of magnitude when considering families of customizable filters and/or filter sets of size ≥ 4. Hence, an alternative to exhaustive search must be formulated to cope with the general case, a goal set forth in this paper.

To pursue this goal, the following sections first overview the two main existing HS-from-RGB reconstruction methods and demonstrate the impact of filter selection on their performance. We then describe strategies for reducing the amount of computations necessary to estimate expected performance of a filter set, as well as a methodology for quickly selecting a close-to-optimal filter set from a very large filter space. Results are shown on several commercial filter spaces.

4. Reconstruction methodology

Our focus is on filter selection for the reconstruction of HS signals from RGB or multispectral cameras, and as we clarify below, this process can be applied to any reconstruction method. Two such methods have been recently suggested by both Nguyen *et al*. [20] as well as Arad and Ben-Shahar [4]. The former proposed a neural network-based approach which produced estimates for both reflectance and scene illumination, while the latter proposed reconstruction

[1] Several manufacturers, for example OMEGA Optical[2], can make bandpass filters with a width from 0.2nm and up to several hundred nanometers in the spectral region between 350nm and 2500nm.

via sparse dictionary representations and reconstruct spectral radiance values. In both methods, evaluating the HS reconstruction performance of a filter set requires a training phase and a testing phase. During the training phase, each system is calibrated to reconstruct HS images from RGB images obtained under the target filter set by processing a set of HS training images. Once calibrated, the performance of each system may then be evaluated in three steps. First a set of test images is produced by simulating the spectral projection of fresh HS source images through the candidate filter set. Second, these test images are fed into the reconstruction algorithm to obtain reconstructed HS images. And finally, the reconstructed HS images are compared to their corresponding ground truth HS source images.

In the following sections we elaborate the steps involved in each of the two reconstruction methods and our evaluation methodology. Reported computation times were obtained on a an Intel Core i5-2400 CPU desktop equipped with 24GB of RAM.

4.1. Nguyen *et al*. [20] Method

The approach proposed by Nguyen *et al*. [20] includes components for recovery of both object reflectance and scene illumination. In this paper, we will limit our discussion to their reflectance recovery methodology which can be summarized as follows:

Training

1. Project each training image to RGB via the selected filter set.
2. Perform illumination correction ("white balance") on projected images.
3. Use corrected-RGB and HS pixels pairs to train a radial basis function (RBF) network.

Testing

1. Perform illumination correction ("white balance") on the test RGB image.
2. Use the RGB network produced by the training phase in order to estimate HS reflectance values for each pixel.

When implemented on our benchmark platform, the training process took ≈ 4 minutes per filter set when performed over the $16,400$ training pixels included by the authors alongside their sample code. Once trained, the system required ≈ 2 minutes to reconstruct $2.5 \cdot 10^6$ HS pixels over 31 channels.

The most computationally intensive step in the training process is step 3, i.e., training of the RBF network. As this step relies on HS images projected through the selected filter set as input, it must be repeated for each evaluated filter set. Hence the time required for each filter set evaluation was ≈ 6 minutes.

4.2. Arad and Ben-Shahar [4] Method

The system proposed by Arad and Ben-Shahar [4] reconstructs HS radiance signatures from RGB pixels and can be summarized as follows:

Training

1. Generate a sparse overcomplete HS dictionary from training images via the K-SVD algorithm[3].
2. Project the HS dictionary from step 1 via the selected filter set and obtain a corresponding overcomplete RGB dictionary.

Testing

1. Use the dictionary produced in step 2 of the training process in order to represent each pixel in the test image via Orthogonal Match Pursuit (OMP)[25].
2. Apply dictionary weights computed in step 1, to the dictionary produced in step 1 of the training process in order to estimate HS radiance values for each pixel.

When implemented on our benchmark platform, the training process took $\approx$ 5 minutes when performed over 10^6 training pixels. Once trained, the system required ≈ 32 sec to reconstruct $2.5 \cdot 10^6$ HS pixels over 31 channels.

The most computationally intensive step in the training process is step 1, i.e., the generation of a HS dictionary (while step 2 is negligible in time). As opposed to Nguyen *et al*. [20], however, this step is completely independent of the evaluated filter set and does not have to be repeated for each evaluated filter set. Hence the time required for each filter set evaluation was ≈ 32 seconds.

4.3. Evaluation Methodology

As mentioned above, we evaluated filter sets by comparing reconstructed HS images to their corresponding ground truth HS source images. In all experiments, reconstruction accuracy is reported as *relative root mean square error* (RRMSE) [4]:

$$E = \frac{\sum_{i,c} \frac{\sqrt{\left(P_{gt_{i_c}} - P_{rec_{i_c}}\right)^2}}{P_{gt_{i_c}}}}{|P_{gt}|} \tag{1}$$

where $P_{gt_{i_c}}$ and $P_{rec_{i_c}}$ denote the c spectral channel value of the i-th pixel in the ground truth and reconstructed images respectively, and $|P_{gt}|$ is the size of the ground truth

image (total pixels $\times$ channels). We emphasize that measuring relative errors (as in [4]), rather than absolute errors (as in [20]) is a much more conservative approach that ensures that errors in low luminance pixels and spectral channels are given equal weight to those in high luminance pixels/channels (even though they are typically much smaller in absolute values).

Experiments utilizing the Nguyen *et al*. [20] method were based on code provided by the authors which was used without modification. Experiments utilizing the Arad and Ben-Shahar [4] method were based on our own implementation, as no sample code has been released by the authors to date. In this implementation, images projected to the camera-domain were truncated to 3 significant digits before further processing in order to simulate the limited dynamic range of camera sensors.

We emphasize again that it was not our intention to compare between the two reconstruction methods or argue that one is superior to the other. This would have been difficult to do, if nothing else then for the different type of output they generate (reflectance + illumination vs. radiance). Instead, our goal is to show that filter selection has a significant impact on reconstruction performance even when estimation methods and target scenes differ.

Finally, we note that although processing speed may not be a primary concern for practical HS-from-RGB systems that act upon a single image at a time, one of the goals of this paper is to devise filter optimization for very large filter spaces. Hence while Sec. 5 shows that both methodologies are significantly impacted by filter selection, our larger scale experiments employ only the Arad and Ben-Shahar [4] reconstruction method, whose processing time per filter set is an order of magnitude smaller.

5. Filter Set Impact on Hyperspectral Estimation

While Oh *et al*. [21] demonstrated that individual camera responses vary enough to allow extrapolations of HS information from a scene imaged by multiple RGB cameras, it does not immediately follow that some of these relatively similar response functions may offer significant gains in the task of single-shot HS recovery from RGB. In order to evaluate the effect of small changes in response functions, we have conducted an experiment evaluating HS reconstruction performance over 28 simulated cameras whose responses were measured by Jiang *et al*. [13].

5.1. Nguyen *et al*. [20] Method

In order to evaluate the impact of filter selection on the Nguyen *et al*. [20] method, RBF networks were trained for each of the 28 camera response functions using unmodified code provided by the authors. The training image set comprised of $16,400$ HS pixels while the test image set comprised of 25 full test images containing a total of $5.6 \cdot 10^7$ HS pixels. Both sets were provided by the authors alongside their code. Fig. 2 depicts sorted RRMSE values of ground truth vs. estimated reflectance across all cameras. While differences in performance between individual cameras may be small, a clear gap and 18.7% improvement are easily observed between the worst (RRMSE=0.293) and the best (RRMSE=0.238) preforming cameras.

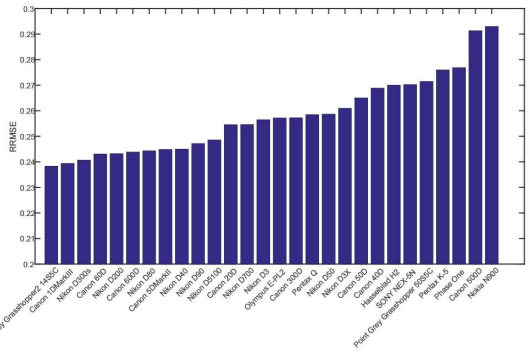


Figure 2: Relative root mean square error (RRMSE) of HS reconstruction per camera response function over all test files using the Nguyen *et al*. [20] method. Note the performance gap between the best and worst performing cameras.

5.2. Arad and Ben-Shahar [4] Method

A similar evaluation process was performed for the Arad and Ben-Shahar [4] method. As most of the training process in this method is independent of the evaluated filter set, the training set comprised of 10^6 pixels randomly selected from 51 BGU HS database[4] training images ($20,000$ pixels per image). A second set of 51 images from the same database was used for the test phase. Fig. 3 describes the reconstruction performance and again, a clear gap and 20.6% improvement are easily observed between the worst (RRMSE=0.160) and the best (RRMSE=0.127) preforming cameras.

In order to allow rapid filter selection from much larger filter spaces, a second experiment was performed over just $2.5 \cdot 10^6$ pixels randomly sampled from the 51 image test set. As Fig. 3 shows, this sampling approach provides virtually identical results. Indeed, the random pixel sample behaves as a very strong predictor for performance over the entire test set (maximum relative difference $< 0.09\%$), thereby facilitating similar evaluation and filter selection over the much larger filter sets to come (Sec. 7).

6. Evolutionary optimization of filter sets

In the previous section we found that HS reconstruction performance may vary significantly across filter sets, even if

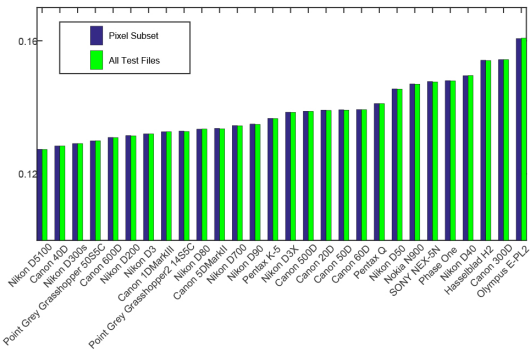


Figure 3: Relative root mean square error (RRMSE) of HS reconstruction per camera response function over all test files (green) and a random selection of pixels from all test files (blue) using the Arad and Ben-Shahar method. Performance over a random pixel subset is a strong predictor of overall performance, with subset RRMSE varying by at most 0.0001 from test set RRMSE. Note the performance gap between the best and worst performing cameras.

only a small number of consumer camera filter sets are considered. Since none of these cameras were designed with HS reconstruction in mind, optimal filter selection over filter spaces beyond such commercial cameras may offer significant performance gains for HS reconstruction.

That said, evaluating performance of a specific filter set in the task of HS estimation may be a computationally intensive task. Depending on the methodology used, reconstruction of a 1MP test image may take hours [15] or seconds [4]. Even while using the latter methodology, exhaustive evaluation of large filter spaces remains impractical.

To further compound the problem, the filter space (the space of filter combinations) is neither continuous nor convex. Hence it is unlikely that convex optimization methods will converge to an optimal solution. Fortunately, evolutionary algorithms are well suited for such problems and may provide a near-optimal solution at a significantly lower computational cost than exhaustive evaluation.

In order to efficiently discover filter sets well-suited for HS reconstruction within large filter spaces, the following evolutionary approach is applied: a small "population" of random filter combinations is generated. These filter sets are evaluated and ranked for reconstruction performance. Finally, a new "population" of filters is generated via "survival of the fittest", crossover, mutation and random replacement:

- **"Survival of the fittest"** preserves the highest rated filter sets.
- **Crossover** combines two filter sets to produce an "offspring" containing randomly selected filters from each of the two sets.
- **Mutation** produces a new filter set by randomly replacing one out of the n filters in a filter set of size n.
- **Random replacement** new filter sets are produced by randomly selecting filters from the entire filter space.

Filter sets from the current population were randomly selected to participate in crossover/mutation with a probability proportional to their fitness ranking. Filter sets for random replacement were selected with equal probability from the entire filter space.

In our experiment each new "population" was composed of: 10% previous populations filters selected by "Survival of the fittest", 40% crossover products of filter sets from the previous population, 10% mutations of filter sets from the previous population and 40% randomly generated filter sets.

This evolutionary optimization process is repeated until the best reconstruction performance remains constant across several iterations, or a predefined run time limit is reached. Note again that whenever possible we carry out both the evolutionary optimization as well as an exshaustive search in order to ratify the former approach through the latter.

7. Evaluation and Results

In order to evaluate the performance of our proposed methodology we have performed evolutionary optimization of 3 filter sets over the following filter spaces:

- A set of 21 filters provided by Midopt[1].
- The set of 84 commercial camera filters measured by Jiang *et al*. [13] (3-filters per camera $\times$ 28 cameras).
- A set of 1022 filters provided by OMEGA Optical [2].

Note that these spaces contain 1330, 95284, and $1.77 \cdot 10^8$ possible 3-filter combinations, respectively.

As detailed in Sec. 4, the Arad and Ben-Shahar [4] method has a significant computational advantage over the Nguyen *et al*. [20] method when evaluation of multiple filter sets is required. Furthermore, since the Nguyen *et al*. [20] method relies on an illumination correction ("white balancing") step, its performance may be adversely impacted by filter sets that do not approximate RGB/CIE-1913. Hence, the following experiments employed only the Arad and Ben-Shahar [4] methodology. The training set in each experiment comprised of 10^6 pixels randomly selected from 51 BGU HS database[4] training images ($20,000$ pixels per image). To further reduce computational cost, the test set comprised of the $2.5 \cdot 10^6$ random pixel sample described in Sec. 5.2. As discussed in Fig. 3, this approach provide highly reliable approximation for a fraction of the computational effort.

7.1. Midopt Filter Space

In order to explore alternatives to consumer cameras, as well as the convergence of evolutionary optimization, a small scale experiment was performed over the Midopt filter space. the Midopt filter collection [1] contains 21 band-pass and "color correction" filters that result in a relatively small $\left(\binom{21}{3} = 1330\right)$ filter space. Exhaustive evaluation was performed, resulting in the RRMSE performance depicted in Fig. 4. Similarly to the results over 28 consumer cameras described in Fig. 3, there is a significant performance gap between the best and worst performing filter sets. Additionally, the best performing filter set provides an average RRMSE of 0.116 - an improvement over consumer cameras (best RRMSE=0.127). Hence we conclude that even a small, generic, filter set may provide an advantage relative to consumer camera RGB filters.

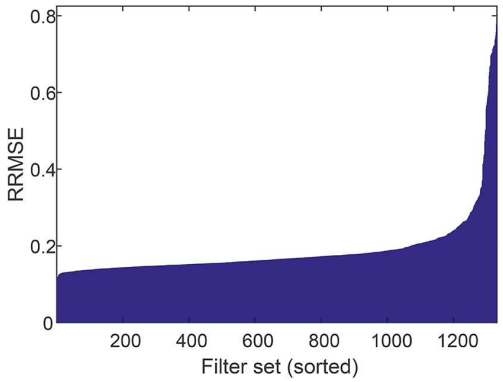


Figure 4: Relative root mean square error of HS reconstruction per filter set in Midopt filter space. Note the performance gap between the best and worst performing filter sets.

While the Midopt filter set may be too small in order to enjoy significant performance gains via evolutionary optimization, a small-scale experiment was performed. Convergence rates were examined by performing 50 repetitions of evolutionary optimization with a population size of 100. On average, optimization converged to the optimal filter set (RRMSE 0.116) within 9.14 iterations (standard deviation 8.07), after examining 499 distinct filters (less than 38% of the entire filter space). Fig. 8a depicts the optimal filter set within the Midopt filter space.

Finally, in order to further verify the representation power of the random pixels subset used during evolutionary optimization, RRMSE values of the top-performing filter set was recomputed over all test images ($9.2 \cdot 10^7$ hyperspectral pixels), producing RRMSE=0.116 yet again.

7.2. Commercial Camera Filter Space

Exploring combinations of band-pass and "color correction" filters provided a significant advantage in HS reconstruction accuracy over existing camera designs. Rather than forgo RGB-like filters completely, it may be beneficial to examine alternative combinations of filters from consumer RGB cameras. Therefore, we consider the filter space defined by all 84 consumer camera filters measured by Jiang *et al*. [13] (three filters per camera $\times$ 28 cameras). These filters span a filter space of size $\binom{84}{3} = 95284$ over which both exhaustive evaluation and evolutionary optimization were performed.

Fig. 5 depicts the RRMSE of HS reconstruction per filter set. The optimal filter set (fig. 8b) included the blue filter from a Canon 5DMarkII, the green filter from a Nikon D40, and the red filter from a Hasselblad H2. Combined as a filter set, they provided a RRMSE of 0.119 - an improvement over each individual camera (Canon 5DMarkII RRMSE 0.133, Nikon D40 RRMSE 0.149, Hasselblad H2 0.154) as well as over the the highest performing consumer camera (Nikon D5100 RRMSE 0.127).

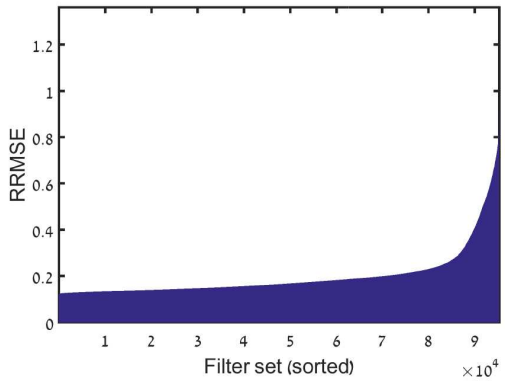


Figure 5: Relative root mean square error of HS reconstruction per filter set in commercial camera filter space. Note the performance gap between the best and worst performing filter sets.

While exhaustive evaluation of 95284 filter sets was required in order to find the globally optimal filter set, evolutionary optimization provided a significantly faster convergence. With a population size of 300, a near optimal filter set (RRMSE 0.120) was found within 4 generations. The globally optimal filter set (RRMSE 0.119) was located after 46 generations, requiring the evaluation of less than 13800 filter sets, namely less than 15% of the entire filter space and the effort done in the exhaustive search. Repeated experiments with the same parameters (50 repetitions) show that the optimal filter set is found, on average, after 34.8 iterations (standard deviation 23.1), examining 8232 distinct filter sets (less than 9% of the entire filter space).

7.3. OMEGA Optical Filter Space

Previous experiments not only demonstrated the utility of non-RGB-like filters in HS reconstruction but also the efficient convergence of evolutionary optimization over large filter spaces. Therefore a final experiment was performed over a large and highly varied filter space. The OMEGA filter space is spanned by 1022 assorted optical filters [2], providing $\binom{1022}{3} = 177388540$ possible combinations. Since exhaustive evaluation of the entire OMEGA Optical filter space would take many processor-years, only evolutionary optimization and random search were performed.

Using a population size of 300, three evolutionary optimizations were repeated, each with a maximal runtime of 50 generations that together evaluated a total of 39,595 filter sets and produced a filter set with a RRMSE of 0.107. The convergence was very fast and obtained a filter set with RRMSE=0.109 after only 4031 evaluations, while the final gain of 0.002 in RRMSE was obtained during the rest of the evolutionary optimization process. Perhaps not surprisingly given the size of the filter set, this result outperforms the best sets obtained from the smaller Midopt and commercial camera filter sets.

To obtain some measure of performance gain over exhaustive search, should we have been able to perform one, we carried out a random search by repeatedly sampling the filter space for sets of size 3 and evaluating their merit for HS reconstruction. This procedure was repeated 450,000 times and produced a filter set with a RRMSE of 0.107. Exhaustive search required 344,600 filter evaluations, an order of magnitude more than evolutionary optimization, to first encounter a filter set with RRMSE=0.107. Figs. 6 and 7 present convergence rates of both processes and Fig. 8c depicts the optimal filter set found within the OMEGA filter space. As in Sec. 7.2, recomputing the RRMSE values of the top performing filter set over all test images ($9.2 \cdot 10^7$ hyper-spectral pixels) produced identical RRMSE values.

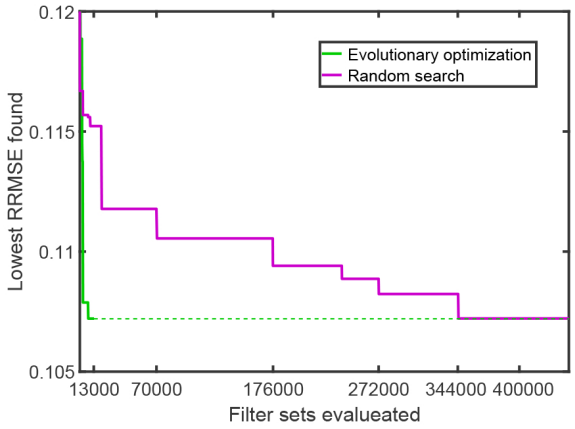


Figure 6: Convergence rates of the best evolutionary optimization (green, cut at after 50 generations) and random search (magenta). Dashed green line denotes the lowest RRMSE found. Note the fast convergence of the optimization.

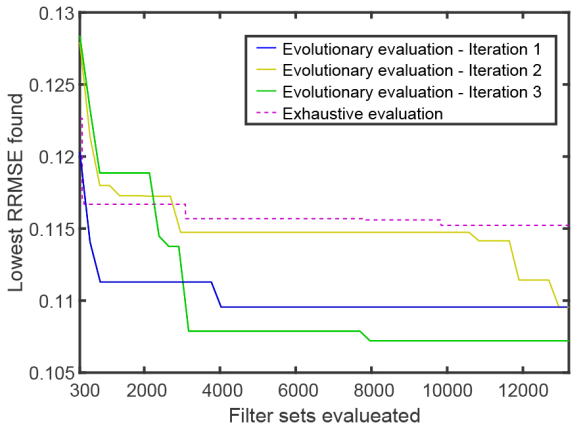
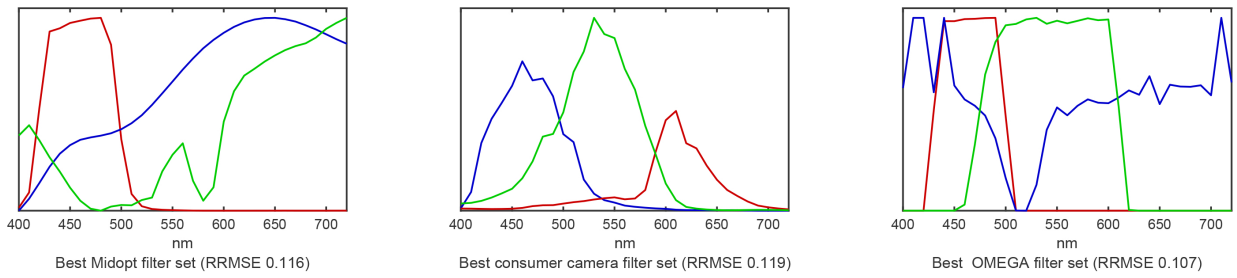


Figure 7: Comparison of convergence rates between all evolutionary iterations (blue, yellow, green) compared to exhaustive evaluation over the same amount of filter set evaluations. Note that all evolutionary optimization attempts converged to RRMSE$\leq$ 0.109 within at most $12,920$ filter evaluations, while exhaustive evaluation required over $175,900$ evaluation to reach the same result and over $238,700$ evaluations to improve upon it.

8. Discussion

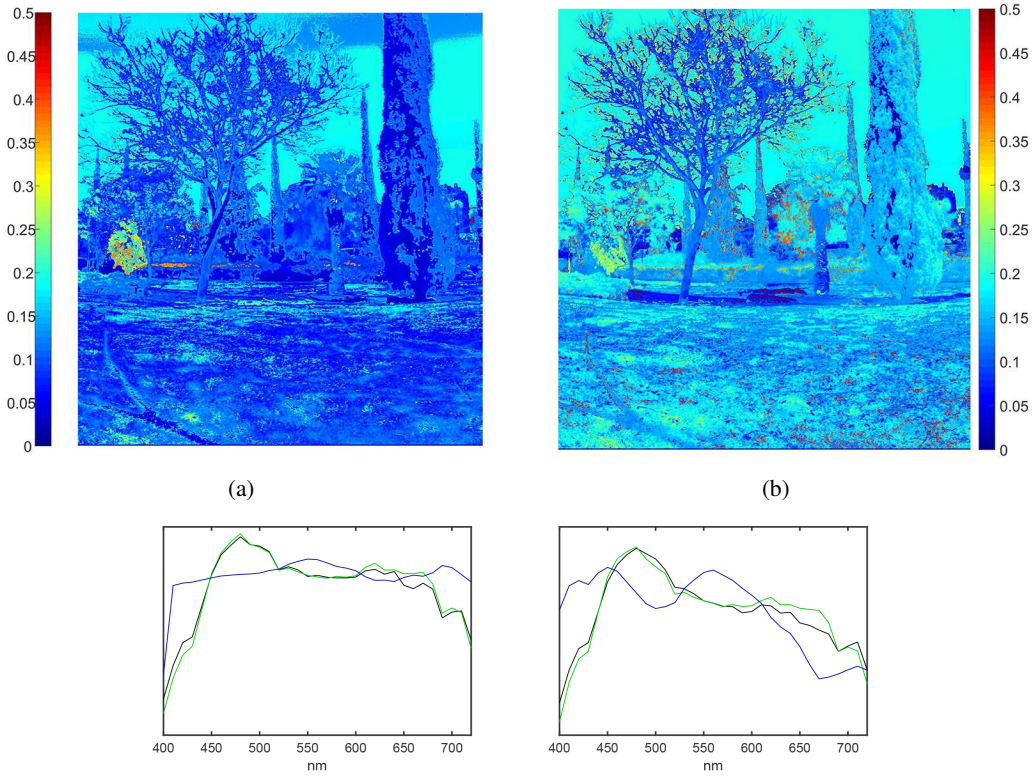
By exploring various filter spaces, we find that hyperspectral estimation errors can be reduced at least 20.6% by simply selecting an optimal consumer camera (cf. Olymus E-PL2 vs. Nikon D5100 when using the Arad and Ben-Shahar [4] method). If custom filter sets are considered, estimation errors can be reduced at least 33.1% (cf. Olymus E-PL2 vs. optimal filter set found in OMEGA filter space). Larger and more varied filter spaces may even provide additional gains. Fig. 9 demonstrates the performance gap between reconstruction using consumer camera filter sets and optimized filter sets. In a field where "state of the art" methods often compete for single percentage point advantages, such a performance gap should not be overlooked. Furthermore, we demonstrated that evolutionary optimization can select optimal or near-optimal filter sets from large filter spaces, while exploring only a small fraction of the filter space and a small sample of the test data.

As our experiments found filter selection to significantly impact performance of two HS reconstruction systems based on different methodologies, there is reason to believe that future HS-from-RGB systems, or even current hybrid RGB-HS systems, can be similarly affected. The de-



(a) (b) (c)

Figure 8: Filter sets with the lowest RRMSE found for Midopt filter space 8a, commercial camera filter space 8b, and OMEGA optical filter space 8c.



(a) (b)

(c)

Figure 9: (a) Average channel RRMSE of a HS image reconstructed using the optimal OMEGA filter set. (b) average channel RRMSE of a HS image reconstructed using the Olymus E-PL2 filter set. (c) Two representative failure cases where spectra reconstructed using the optimal OMEGA filter set (green) match ground truth spectra from the test pixel subset (black) while spectra reconstructed using the Olymus E-PL2 filter set (blue) did not.

sign of any such system should therefore take filter selection into account and may employ our suggested methodology for optimizating this process.

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.

References

[1] MIDOPT website. http://midopt.com/. Accessed: 2016-05-16.

[2] OMEGA Optical. http://www.omegafilters.com/. Accessed: 2016-05-16.

[3] M. Aharon, M. Elad, and A. Bruckstein. K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation. *IEEE TRANSACTIONS ON SIGNAL PROCESSING*, 2006.

[4] B. Arad and O. Ben-Shahar. Sparse Recovery of Hyperspectral Signal from Natural RGB Images. *European Conference on Computer Vision*, pages 19–34, 2016.

[5] X. Cao, X. Tong, Q. Dai, and S. Lin. High Resolution Multispectral Video Capture with a Hybrid Camera System. In *CVPR*, 2011.

[6] C. Chi, H. Yoo, and M. Ben-Ezra. Multi-spectral imaging by optimized wide band illumination. *International Journal of Computer Vision*, 86(2):140, Nov 2008.

[7] M. Descour and E. Dereniak. Computed-tomography imaging spectrometer: experimental calibration and reconstruction results. *Applied Optics*, 1995.

[8] H. Du, X. Tong, X. Cao, and S. Lin. A prism-based system for multispectral video acquisition. *ICCV*, 2009.

[9] D. W. Fletcher-Holmes and A. R. Harvey. Real-time imaging with a hyperspectral fovea. *Journal of Optics A: Pure and Applied Optics*, 7(6):S298–S302, jun 2005.

[10] M. Goel, S. N. Patel, E. Whitmire, A. Mariakakis, T. S. Saponas, N. Joshi, D. Morris, B. Guenter, M. Gavriliu, and G. Borriello. HyperCam. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*, pages 145–156, New York, New York, USA, 2015. ACM Press.

[11] M. D. Grossberg and S. K. Nayar. What is the space of camera response functions? In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 2, pages II–602. IEEE, 2003.

[12] M. Huber-Lerner, O. Hadar, S. R. Rotman, and R. Huber-Shalem. Hyperspectral Band Selection for Anomaly Detection: The Role of Data Gaussianity. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(2):732–743, feb 2016.

[13] J. Jiang, D. Liu, J. Gu, and S. Süsstrunk. What is the space of spectral sensitivity functions for digital color cameras? In *Applications of Computer Vision (WACV), 2013 IEEE Workshop on*, pages 168–179. IEEE, 2013.

[14] W. R. Johnson, D. W. Wilson, and G. Bearman. Spatial-spectral modulating snapshot hyperspectral imager. *Applied optics*, 2006.

[15] R. Kawakami, J. Wright, T. Yu-Wing, Y. Matsushita, M. Ben-Ezra, and K. Ikeuchi. High-resolution hyperspectral imaging via matrix factorization. *CVPR*, 2011.

[16] H. Kwon and Y. W. Tai. RGB-guided hyperspectral image upsampling. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 11-18-Dece, pages 307–315, 2016.

[17] J.-P. Ma, Z.-B. Zheng, Q.-X. Tong, and L.-F. Zheng. An application of genetic algorithms on band selection for hyperspectral image classification. In *Proceedings of the 2003 International Conference on Machine Learning and Cybernetics (IEEE Cat. No.03EX693)*, volume 5, pages 2810–2813 Vol.5, Nov 2003.

[18] J. Minet, J. Taboury, F. Goudail, M. Péalat, N. Roux, J. Lonnoy, and Y. Ferrec. Influence of band selection and target estimation error on the performance of the matched filter in hyperspectral imaging. *Applied Optics*, 50(22):4276, aug 2011.

[19] J. Nakamura. *Image sensors and signal processing for digital still cameras*. CRC press, 2016.

[20] R. M. H. Nguyen, D. K. Prasad, and M. S. Brown. Training-based spectral reconstruction from a single RGB image. In *ECCV*, volume 8695 LNCS, pages 186–201. Springer International Publishing, 2014.

[21] S. W. Oh, M. S. Brown, and M. Pollefeys. Do It Yourself Hyperspectral Imaging with Everyday Digital Cameras. *2016 IEEE Conf. Comput. Vis. Pattern Recognit.*, pages 2461–2469, 2016.

[22] T. Okamoto and I. Yamaguchi. Simultaneous acquisition of spectral image information. *Optics Letters*, 1991.

[23] J. I. Park, M. H. Lee, M. D. Grossberg, and S. K. Nayar. Multispectral imaging using multiplexed illumination. 2007.

[24] M. Parmar, S. Lansel, and B. A. Wandell. Spatio-spectral reconstruction of the multispectral datacube using sparse recovery. 2008.

[25] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In *Signals, Systems and Computers, 1993. 1993 CONFERENCE Record of The Twenty-Seventh Asilomar CONFERENCE on*, 1993.

[26] T. Wang, Z. Zhu, and H. Rhody. A smart sensor with hyperspectral/range fovea and panoramic peripheral view. *CVPR*, 2009.

[27] G. Zhu, Y. Huang, J. Lei, Z. Bi, and F. Xu. Unsupervised Hyperspectral Band Selection by Dominant Set Extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 54(1):227–239, jan 2016.