

Illumination Compensation and Normalization Using Low-Rank Decomposition of Multispectral Images in Dermatology

Alexandru Dului¹(✉), Richard Brosig¹, Saahil Ognawala¹,
Tobias Lasser¹, Mahzad Ziai², and Nassir Navab¹

¹ Computer Aided Medical Procedures (CAMP),
Technische Universität München, Munich, Germany
dului@in.tum.de

² Department of Dermatology and Allergology,
Klinikum Rechts der Isar, Munich, Germany

Abstract. When attempting to recover the surface color from an image, modelling the illumination contribution per-pixel is essential. In this work we present a novel approach for illumination compensation using multispectral image data. This is done by means of a low-rank decomposition of representative spectral bands with prior knowledge of the reflectance spectra of the imaged surface. Experimental results on synthetic data, as well as on images of real lesions acquired at the university clinic, show that the proposed method significantly improves the contrast between the lesion and the background.

1 Introduction

When acquiring images with a color camera, spatial distribution of the incoming illumination and geometry of an object have a great impact on image formation. In applications, such as lesion classification in dermatology, the actual color (diffuse spectral reflectance or albedo) of the tissue is the main point of interest, and not how its surface has interacted with the illumination. As intensity values of the image pixels only store how the surface has interacted with the incoming illumination, we are looking for a way to separate the albedo A from the illumination L in an image I :

$$I = A \odot L \tag{1}$$

where I , A are multi-channel images, L is a single-channel image (same size as I and A), and \odot is the component-wise product of the pixel values.

Multispectral imaging is a powerful tool for tissue classification, due to its improved spectral resolution compared to conventional RGB imaging, with the latter only approximating the subjective color perception of the human eye. This improved spectral resolution allows for better tissue discrimination, however this is hampered by the presence of large illumination changes across the image. As lesion classification methods [3] rely on many features that can be observed on a lesion, such as color and fine structures, they are quite sensitive to illumination

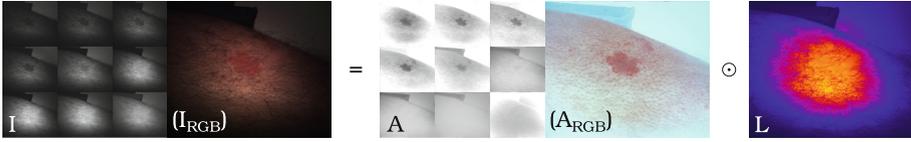


Fig. 1. Example of separating a multispectral image I into albedo A and illumination L (false-colors) using our method. I_{RGB} and A_{RGB} are projections of their respective multispectral images, I and A , onto RGB space. Here and in all following figures, a heat color-map is used to visualize the illumination L (Color figure online).

changes. Our motivation is to exploit the additional information of multispectral imaging, in order to compensate for such illumination changes, while preserving both morphological and colorimetric (spectral) features of the lesions.

In applications that recover the albedo or just compensate for illumination for detail enhancement [1, 10], it is common practice to acquire multiple images while keeping the viewpoint fixed and varying the illumination. The assumption is made that the albedo remains the same while the illumination changes. In our work we also acquire multiple images under a fixed viewpoint, however we do not vary the illumination but the spectral range for each image channel. We assume that it is the spatial distribution of the illumination that remains constant and that the albedo varies.

We introduce a two step approach based on low-rank decomposition of the multispectral image I in order to approximate a per-pixel illumination map L which enables the recovery of the diffuse albedo A . We evaluate the method on multispectral images of synthetic lesions on volunteers and of real lesions of patients.

2 Related Work

When considering Eq. (1), the image I cannot be separated into A and L without additional information.

Photometric stereo methods [10] acquire multiple images, while keeping the viewpoint fixed. In each image, the scene is illuminated from a different calibrated direction. Here, both the the surface geometry and albedo are estimated together. More recently [1], the requirement for calibrated illumination has been relaxed, only assuming the light sources to be sufficiently distant. Although multispectral imaging can be adapted to such multi-light acquisitions [9] (which considerably increases acquisition times), our focus is on illumination compensation from a single multispectral image.

Similar data as in [1] (fixed viewport multi-light images) is used by image enhancement methods [6, 14] which, rather than removing the illumination altogether, focus on minimizing its influence, while preserving fine features, that would otherwise be lost without illumination cues. Fattal *et al.* [6] propose an image decomposition method for multi-light image sequences that removes large illumination artefacts (like shadows) while enhancing small surface details due

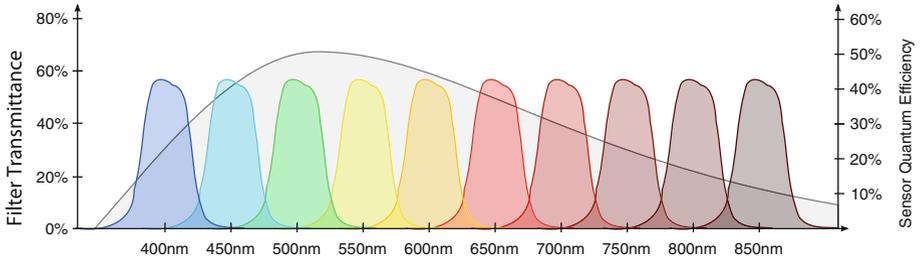


Fig. 2. Example transmission spectra of band-pass filters (colored functions), and example spectral sensitivity of image sensor (larger gray function) (Color figure online).

to shading. Such methods have a simpler model than photometric stereo of how light interacts with the scene as they only focus on image enhancement.

These methods [1, 6, 10, 14] all exploit variations in illumination in order to recover the albedo. This is not unlike our method, we however make the reverse assumption, that the albedo varies and that the illumination stays the same.

Influences of illumination can also be modeled in single images when enough assumptions can be made about the scene. Shi *et al.* [11] propose an adaptive linear function that estimates the background of historical documents, compensating for uneven shading due to paper geometry and non-uniform illumination. For robust face recognition, Chen *et al.* [2] propose an illumination normalization approach based on a Discrete Cosine Transform in the logarithmic domain. Although not exactly modeling illumination, the methods focuses on preserving application relevant features while removing most image illumination.

Vogel *et al.* [13] present a radiometrically calibrated multispectral imaging system to assess tissue vasculature. The authors model both the distribution of the illumination across the image and the geometry of the subject. The former is mostly corrected for by the radiometric calibration of the camera and the calibrated illumination, leaving only the influence of the geometry. Assumptions are made about the geometry of the object and a simple curvature model is fitted to the data in order to compensate for the remaining shading. This method has the most similar input data and hardware to our own, however, most of the influences of illumination are corrected through a calibrated and controlled acquisition environment. Unlike Vogel *et al.* [13] we make no assumptions about either the distribution of light intensity or the geometry of the scene.

3 Methods

Our method exploits the increased spectral resolution of multispectral images in order to distinguish between the most important spectra that influence image formation: illumination, background and foreground. As this method was initially developed for dermatology, the background is the skin and the foreground is the lesion we are investigating. The method can also be used on multiple foreground and/or background spectra as long as certain criteria are met (see Sect. 3.2).

3.1 Multispectral Image Formation

For simplicity we will focus only on the spectral information and how that relates to pixel intensity values without considering complete imaging optics.

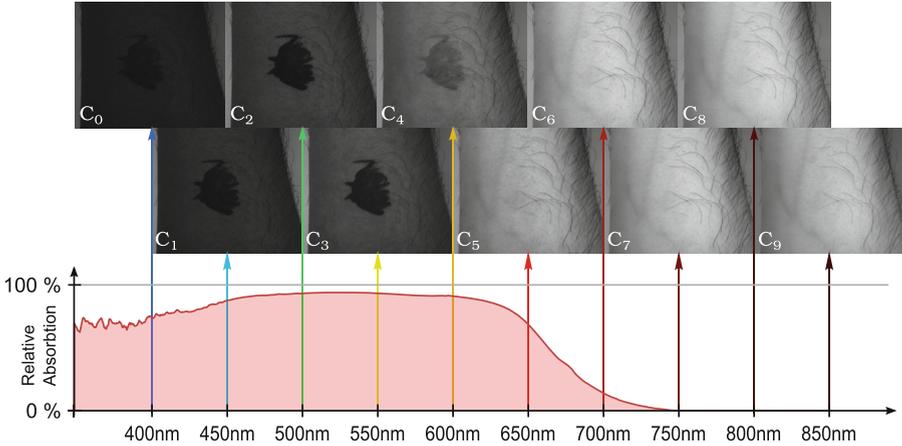


Fig. 3. Example multispectral image of a forearm marked with a red dye that has a similar absorption spectrum as an erythemous lesion (Fig. 1). Above 700 nm the dye is no longer distinguishable from the skin, just as its relative absorption spectrum would indicate (relative to skin) (Color figure online).

The basic layout of a multispectral camera is analogous to that of most RGB camera: lens, filters and gray-scale sensor. Where a RGB camera has a *color filter array* (CFA) consisting of tiny red, green and blue filters mounted directly on the sensor, a multispectral camera can have different designs in order to support a considerably larger number of filters. Although we will only focus on filter-wheel cameras¹, our method can be applied on images acquired with any type multispectral camera.

Sensors used in multispectral cameras have a wide spectral sensitivity (Fig. 2), however they cannot distinguish the wavelength of the photons they detect. The transmission spectra (Fig. 2) of the camera filters act as windowing functions, letting only a certain part of the spectrum through. This is how both RGB and multispectral cameras are able to distinguish between different parts of spectrum of the incoming light: they use the filters to select what wavelengths reach the sensor. Therefore the intensity value $C_{i,p}$ for a pixel p of channel i is:

$$C_{i,p} = \int Sl_p(\lambda) \cdot St_i(\lambda) \cdot Ss(\lambda) d\lambda \tag{2}$$

¹ *Filter-wheel multispectral cameras* sequentially acquires multiple images of different spectral bands, each time exposing the sensor through a different filter.

with Sl_p the spectrum of the light entering the camera, St_i the transmission spectrum of the i^{th} channel, Ss the sensitivity spectrum of the sensor and wavelength λ . Considering that we store the spectra as discrete vectors, Eq. (2) can also be written as:

$$C_{i,p} = (Sl_p \odot St_i) \cdot Ss \tag{3}$$

3.2 Modeling Illumination

Illumination compensation can be viewed as reversing image formation in order to separate the multispectral diffuse reflectance A from the illumination L . Handling a multispectral image $\mathbf{I} \in \mathbb{R}^{m \times n \times k}$ consisting of k channels ($I = (C_i)_{i=1}^k$), can be simplified by linearizing all channels. Therefore a channel $C_i \in \mathbb{R}^{nm}$ can be described by:

$$C_i = A_i \odot (L \cdot f_i) \tag{4}$$

with the global illumination term $f \in \mathbb{R}^k$ and the local illumination term $L \in \mathbb{R}^{nm}$ and a channel and pixel dependant albedo $A_i \in \mathbb{R}^{nm}$. The reshaped I can be modeled as the component-wise multiplication of A with the *outer product* between L and f :

$$I = A \odot (L \cdot f^T) \tag{5}$$

Modeling this problem like this makes several assumptions

The emission spectrum is the same for all light sources. The concept of the channel-wise multiplicative factor f in Eq. (5) holds only for a globally constant emission spectrum. This generally applies to indoor environments, with similar light bulbs and no sunlight. If this were not the case, the illumination would vary spatially across the image depending on the observed channel. It would even be possible to measure this factor beforehand, however since our method is estimating this factor (see Sect. 3.3) this is not necessary.

The bidirectional reflectance distribution function (BRDF) of both background and foreground does not vary with the wavelength. This assumption is key in our

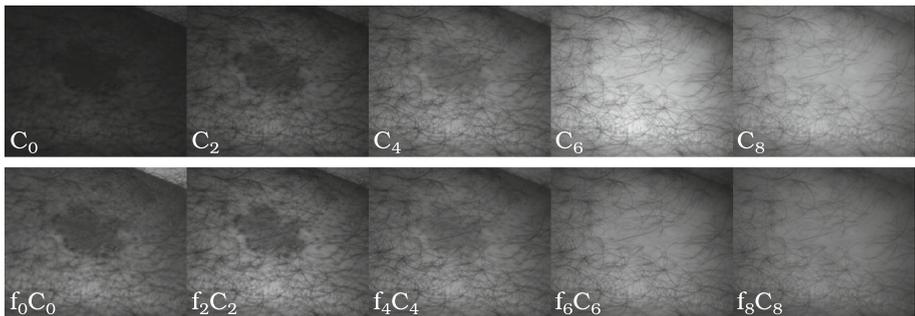


Fig. 4. Selected channels of the multispectral image (detail view of lesion from Fig. 1): C_i before (top row) and $f_i C_i$ after (bottom row) global illumination correction/normalization.

formulation of the local illumination term L , enabling us to consider it the same for all channels and as such the low-rank component.

The foreground spectrum is not completely distinguishable from the background spectrum; this is to say that in certain channels the reflectance of the foreground (lesion) is identical to that of the background (skin), see Fig. 3. We select the channels where this is the case as they only contain the background and illumination. This requires either prior knowledge about the reflectance spectra of both foreground and background or manual selection of the channels. When estimating the local illumination term L , the low-rank decomposition is only performed on these channels.

3.3 Low-Rank Decomposition

We consider I as a set of k channels $C_i \in \mathbb{R}^{mn}$, $i \in \{1, \dots, k\}$ that show the same scene under different acquisition conditions (here, different spectral bands) and compute the correction factors.

We differentiate between two illumination effects. *Global* and *local*. First, the *global* illumination $f \in \mathbb{R}^k$ is a channel wise multiplicative factor correcting global properties like filter permeability or the lamp spectrum. Therefore the set of k channels (C_i) of the same image I are linearly dependent, excluding noise and non-multiplicative foreground denoted as S . Consequently, the matrix $\mathbf{C} \in \mathbb{R}^{mn \times k}$, containing the linearized images as columns, has to have rank one. However, assuming noisy images or foreground structures, which do not have the same spectral response, we are searching for a rank-1 matrix \mathbf{L} with

$$\mathbf{C} = \mathbf{S} + \mathbf{L}, \quad (6)$$

where \mathbf{S} accommodates the non-multiplicative residuals. Based on Eq. (6), the global, multiplicative illumination correction factors f can be recovered by the unique decomposition of

$$\mathbf{L} = l f^T, \quad (7)$$

with $l \in \mathbb{R}^{mn}$, $f \in \mathbb{R}^k$ and $\|f\|_1 = 1$.

In the second step, we calculate the *local* illumination correction. It is based on the same principle, but additionally exploits properties of multispectral acquisition. We take the channels C_i , corresponding to acquisitions in the range over 700 nm, and stack them column-wise in $\tilde{\mathbf{C}}$. Hence, the vectorized *local* illumination $\tilde{l} = L(\cdot)$ can be estimated by

$$\tilde{\mathbf{C}} = \tilde{\mathbf{S}} + \tilde{\mathbf{L}} = \tilde{\mathbf{S}} + \tilde{l} \tilde{f}^T \quad (8)$$

For both matrix decompositions in Eqs. (6) and (8), we use an algorithm similar to the one of Cui *et al.* [4]. Revising formula (6), we want to decompose \mathbf{C} in a rank-1 matrix \mathbf{L} and a matrix \mathbf{S} . The matrix \mathbf{S} models the foreground or noise, and thus should be sparse. This leads to following objective

$$\min \|\mathbf{S}\|_0 \quad \text{s.t.} \quad \mathbf{C} = \mathbf{S} + \mathbf{L} \quad \text{and} \quad \text{rank}(\mathbf{L}) = 1 \quad (9)$$

Similar to [4], we relax $\|\cdot\|_0$ to $\|\cdot\|_{2,0}$ norm, which is a row-wise 2-norm followed by column wise 0-norm. Both are optimized with an greedy algorithm, which alternately optimizes for the rank-1 condition and the sparsity condition. Obtaining a rank-1 approximation is achieved by Singular Value Decomposition (SVD). For the sparsity approximation, we enforce $\|\mathbf{S}\|_{2,0} < \alpha$, by setting $\mathbf{S}_{ij} = 0$ except for the α largest rows in the $\|\cdot\|_2$ norm, therefore α is an upper bound on the number of non-zero rows. The threshold α is roughly set to the number of foreground pixels.

Finally, the albedo A_i is reconstructed using Eq. (4) by replacing L with a reshaped \tilde{l} from Eq. (8) and plugging in the estimate f_i from Eq. (7),

$$A_i = f_i \cdot C_i \oslash L = f_i \cdot C_i \oslash \tilde{\mathbf{I}} \quad \forall i \in \{1, \dots, k\} \quad (10)$$

where \oslash denotes the component-wise division.

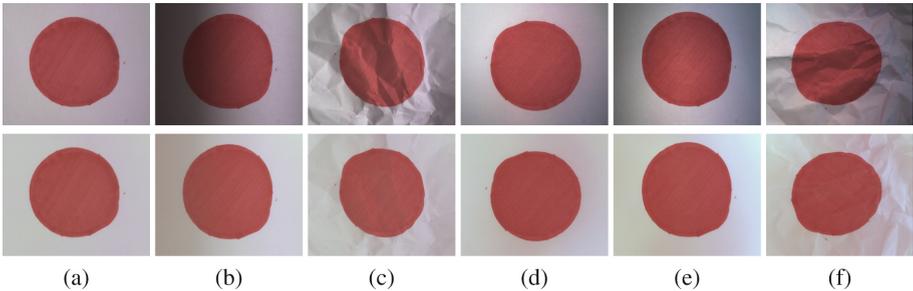


Fig. 5. Red dye on paper under varying illumination conditions: (a,b,c) uniform light from the right, (d,e,f) spot light. The paper was: flat (a,d), bent (b,e) and crumpled (c,f). Top row depicts the original images and the bottom row the corrected images (Color figure online).

4 Experiments

We have acquired multispectral images in order to evaluate our method under multiple scenarios. We start with simple scenarios, where both the scene and the illumination are controlled and move on to images of real lesions of actual patients under varying lighting conditions.

Synthetic Lesions on Paper: Our most simple test scenario involves images of red dye on white paper (Fig. 5). Here we vary both the illumination and the geometry of the paper. The first three images (Fig. 5a–c) are acquired under uniform illumination (same light intensity across the image) from the right, while varying the geometry of the paper: flat (a), bent (b) and crumpled (c). The other three images are acquired under a non-uniform illumination (light is more bright

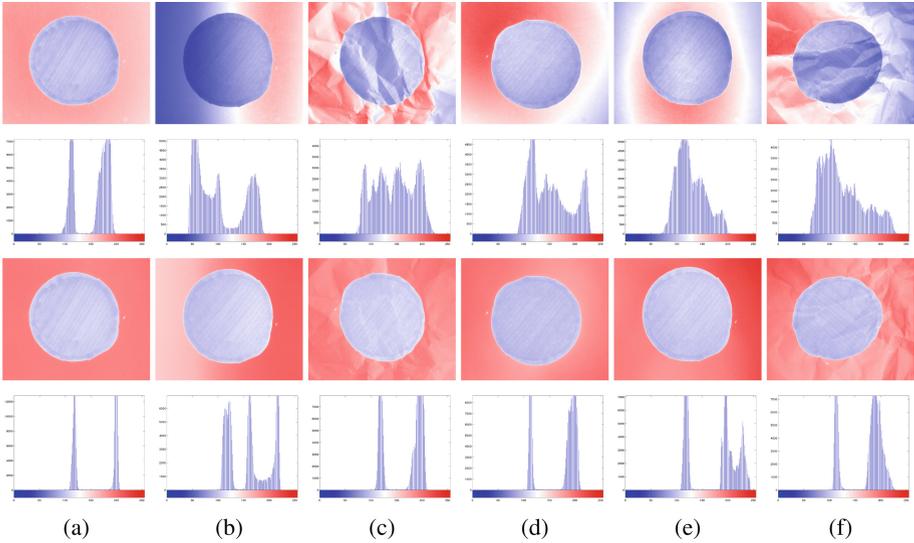


Fig. 6. The grayscale version of the images from Fig. 5 are accentuated with false colors (divergent color-map: blue/white/red) to visualize the drop in image variance after correction (bottom two rows). The histograms also reflect this, clearly showing the two regions (red dye and background) (Color figure online).

in some regions of the image than others), while under the same deformations as the first three: flat (d), bent (e) and crumpled (f). The second row of Fig. 5 shows the same images after applying our method. A considerable improvement in the contrast between foreground and background can be observed, with shading caused by illumination almost completely removed.

Figure 6 shows false color images representing grayscale versions of the same images from Fig. 5: top two rows before and bottom rows after applying our method. Additionally the histogram of each image is also shown in Fig. 6, visualizing how the two areas (red foreground and white background) are much better defined.

Synthetic Lesions on Skin: Images were taken of the forearms of multiple volunteers, with ages ranging from 22 to 34 years (Fig. 7). Red dye was painted on the skin and the images were acquired under uniform illumination, which was provided by multiple halogen lamps behind diffuser curtains (to improve uniformity).

Since the dye was applied uniformly, we can also use this set to validate the performance of our illumination compensation. As can be observed in Fig. 8, the spectra of pixels (both dye and skin) vary considerably less after *local illumination compensation*. It can also be observed that the shape of the spectra, which dictates their color, has not changed. This shows that our method preserves the spectral signature of the tissue.

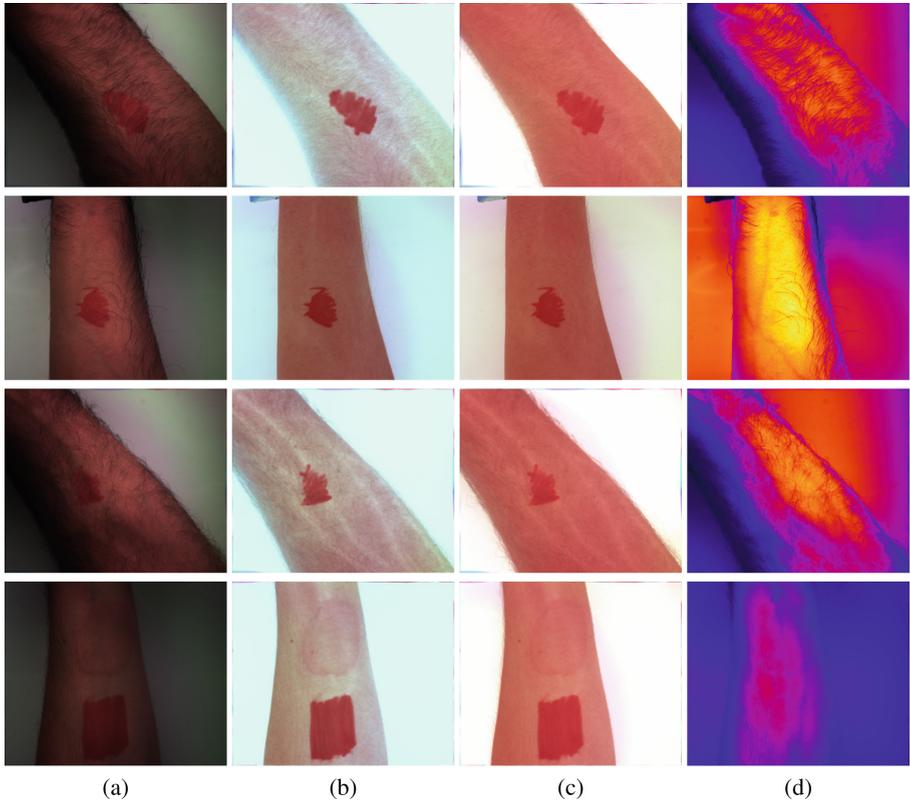


Fig. 7. Each row depicts a synthetic lesion of a different volunteer, with: (a) original image, (b) both local and global illumination compensated, (c) only local illumination compensated without global compensation and (d) illumination map L (false colors).

Psoriasis Lesions: Images were acquired at the university clinic during regular consultation hours of patients with confirmed diagnosis of *psoriasis vulgaris*, with ages between 19 and 74 years. All patients were undergoing treatment with positively responding lesions that were reducing in size. Illumination was provided by a regular halogen examination lamp that had to be repositioned depending on the location of the lesion on the patients body (Fig. 9 column-wise: lower leg, elbow, upper arm and left wrist).

5 Discussion

In Sect. 3.2 we have stated three assumptions essential to our method, with the latter two dictating how we estimate the illumination map L . We will now discuss two special cases, where we observed that these assumptions do not hold completely.

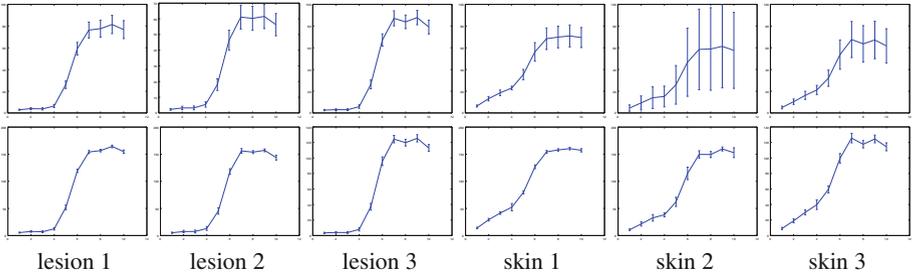


Fig. 8. Compensating for illumination reduces the standard deviation of pixels intensities from similar tissue types. Here we selected lesion and skin pixels from the first three images in Fig. 7. The top row shows the mean spectrum and standard deviation of lesion pixels (left three), and bottom row of skin pixels (right three). The second row shows the same metric for the same pixel sets but, after local illumination correction.

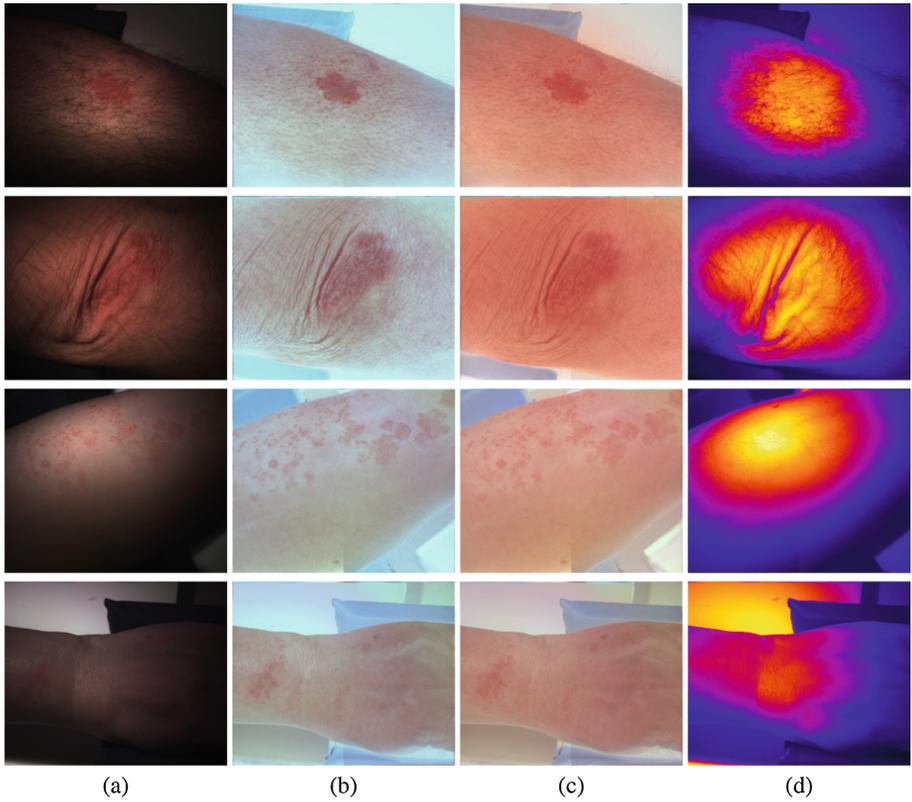


Fig. 9. Each row depicts a different psoriasis lesion(s) (lower leg, elbow, upper arm and left wrist), with: (a) original image, (b) both local and global illumination compensated, (c) only local illumination compensated without global compensation and (d) illumination map L (false colors).

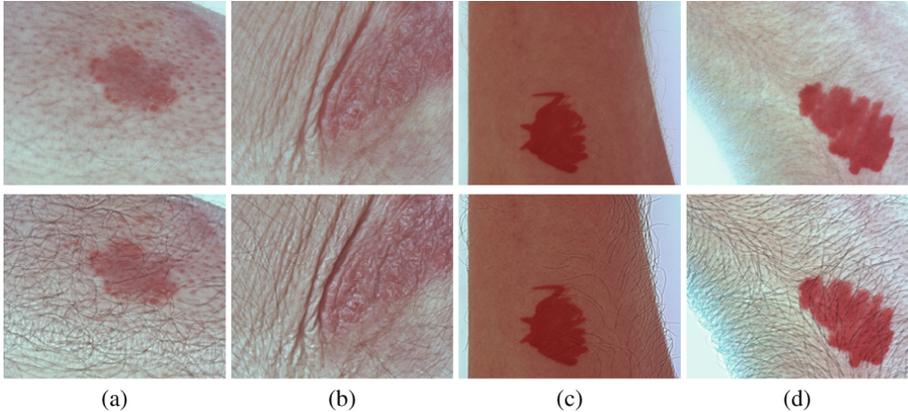


Fig. 10. Normalized albedo of several lesions (psoriasis and red dye) with the top row using unprocessed illumination maps, and the bottom processing the illumination map with a bilateral filter before albedo recovery. Due to their similar size, specular highlights on fine geometric details are introduced and can be observed in the bottom row. Bilateral filtering of L has no impact on the global normalization. These are magnified regions of images shown in Figs. 7 and 9 (Color figure online).

Although we assume the BRDF of both skin and lesions to be constant across the visible spectrum, which it mostly is, we could observe in Fig. 7 that subjects of rows 1 and 3 had clearly visible blood vessels under the skin. Above 750 nm, skin becomes increasingly translucent, potentially revealing underlying blood vessels which are then included in the illumination map L . This introduced bias brightens the areas where the blood vessels are visible, however the spectrum of the tissue is still preserved. We assume two factors to be responsible for this: the relative young age of the subjects in this test, and that all measurements were taken of the inside forearm, where blood vessels are especially prominent. We could not observe this for patients with psoriasis lesions.

We could also observe that the reflectance spectrum of hair does not conform to our assumption (see Sect. 3.2) regarding the distinguishability of the foreground and background spectra. Hair has a spectrum clearly distinguishable from both lesions and skin, and is often included in L . This distinction can be observed in both Figs. 3 and 4, where the hair is clearly black across all images, thus different from both lesion and skin.

Removing the hair from the illumination map would imply to restoring it to the corrected image, and thus the albedo A . Although we could employ a specialized method to remove the hair from L (such as described in [8]), we believe a more generalized approach would be better suited, as it could also encompass fine geometric details.

The task then becomes the removal of local features from L , while preserving the larger scale illumination. We believe this to be similar to the principle of local adaptation, used for tonemapping high dynamic range images [5, 7]. As discussed

in Sect. 2, Fattal *et al.* [6] use a multiscale decomposition based on bilateral filtering to preserve local features.

We employ a simplified implementation [12] with only one level of bilateral filtering, with parameters tuned to preserve larger features in the illumination map, such as skin folds, while removing fine ones, such as hair. Hairs are clearly removed from L and restored to A (see Fig. 10), and although some halo artefacts can be observed, these are considerably less pronounced than if a Gaussian filter were used.

Although L approximates the distribution of illumination across the image, we can observe that treating hair as illumination has the benefit of reducing its visibility in A and further improving image quality.

6 Conclusion

We presented a novel illumination normalization method for use on multispectral images in dermatology to better distinguish between lesions and their background. The spatial distribution of illumination in an image can be recovered by exploiting differences between the spectra of the materials (tissues) being imaged and that of the incoming light. We achieve this by means of a low-rank decomposition that estimates the local variations in illumination as well as global variations across the spectrum. Experimental results on both synthetic and clinical data show significant improvements in image quality.

Acknowledgements. This work was partially funded by the TUM Graduate School of Information Science in Health.

References

1. Basri, R., Jacobs, D., Kemelmacher, I.: Photometric stereo with general, unknown lighting. *Int. J. Comput. Vis.* **72**(3), 239–257 (2007)
2. Chen, W., Er, M.J., Wu, S.: Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **36**(2), 458–466 (2006)
3. Cheng, Y.I., Swamisai, R., Umbaugh, S.E., Moss, R.H., Stoecker, W.V., Teegala, S., Srinivasan, S.K.: Skin lesion classification using relative color features. *Skin Res. Technol.* **14**(1), 53–64 (2008)
4. Cui, X., Huang, J., Zhang, S., Metaxas, D.N.: Background subtraction using low rank and group sparsity constraints. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) *ECCV 2012, Part I. LNCS*, vol. 7572, pp. 612–625. Springer, Heidelberg (2012)
5. Durand, F., Dorsey, J.: Fast bilateral filtering for the display of high-dynamic-range images. *ACM Trans. Graph. (TOG)*. **21**, 257–266 (2002)
6. Fattal, R., Agrawala, M., Rusinkiewicz, S.: Multiscale shape and detail enhancement from multi-light image collections. *ACM Trans. Graph.* **26**(3), 51 (2007)
7. Li, Y., Sharan, L., Adelson, E.H.: Compressing and companding high dynamic range images with subband architectures. *ACM Trans. Graph. (TOG)*. **24**, 836–844 (2005)

8. Nguyen, N.H., Lee, T.K., Atkins, M.S.: Segmentation of light and dark hair in dermoscopic images: a hybrid approach using a universal kernel. In: SPIE Medical Imaging. p. 76234N, International Society for Optics and Photonics (2010)
9. Park, J.I., Lee, M.H., Grossberg, M.D., Nayar, S.K.: Multispectral imaging using multiplexed illumination. In: IEEE 11th International Conference on Computer Vision, ICCV 2007, pp. 1–8 (2007)
10. Samaras, D., Metaxas, D., Fua, P., Leclerc, Y.G.: Variable albedo surface reconstruction from stereo and shape from shading. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. vol. 1, pp. 480–487. IEEE (2000)
11. Shi, Z., Govindaraju, V.: Historical document image enhancement using background light intensity normalization. In: Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004, vol. 1, pp. 473–476. IEEE (2004)
12. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: Sixth International Conference on Computer Vision, 1998, pp. 839–846. IEEE (1998)
13. Vogel, A., Chernomordik, V.V., Demos, S.G., Pursley, R., Little, R.F., Tao, Y., Gandjbakhche, A.H., Yarchoan, R., Riley, J.D., Hassan, M., et al.: Using noninvasive multispectral imaging to quantitatively assess tissue vasculature. *J. Biomed. Opt.* **12**(5), 051604–051604 (2007)
14. Zheng, J., Li, Z., Rahardja, S., Yao, S., Yao, W.: Collaborative image processing algorithm for detail refinement and enhancement via multi-light images. In: IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), pp. 1382–1385. IEEE (2010)