A Customized Camera Imaging Pipeline for Dermatological Imaging

Hakki Can Karaimer[1]   Iman Khodadad[2]   Farnoud Kazemzadeh[2]   Michael S. Brown[1]

[1]York University, Toronto   [2]Elucid Labs

{karaimer, mbrown}@eecs.yorku.ca   {iman, farnoud}@elucidlabs.ca

Abstract

This paper describes the customization of the camera processing pipeline of a machine vision camera that has been integrated into a hand-held dermatological imaging device. The device uses a combination of visible and non-visible spectral LEDs to allow capture of visible RGB imagery as well as selected non-visible wavelengths. Our customization involves two components. The first component is a color calibration procedure that ensures the captured images are colorimetrically more accurate than those obtained through the machine vision camera's native API. The need for color calibration is a critical component that is often overlooked or poorly understood by computer vision engineers. Our second component is a fast method to integrate the narrow band spectral images (some of which are outside the visible range) into the visible RGB image for enhanced visualization. This component of our pipeline involves evaluating several algorithms capable of multiple image fusion to determine the most suitable one for our application. Quantitative and subject results, including feedback from clinicians, demonstrate the effectiveness of our customization procedure.

1. Introduction

This paper describes a custom imaging device for dermatological inspection constructed with an integrated machine vision camera. Machine vision cameras offer several advantages over consumer-oriented cameras, such as DSLR and mobile phone cameras. These advantages include: (1) a larger sensor offering improved signal-to-noise performance; (2) the lack of an near-infrared (NIR) filter that allows non-visible spectral data to be captured; and (3) supporting software and APIs that allow low-level control over camera settings and the ability to perform external event triggering for image capture. One crucial disadvantage of machine vision cameras, often overlooked by computer vision researchers, is that these cameras are typically not colorimetrically calibrated. Moreover, their supporting APIs often do not provide any mechanism for color correc-
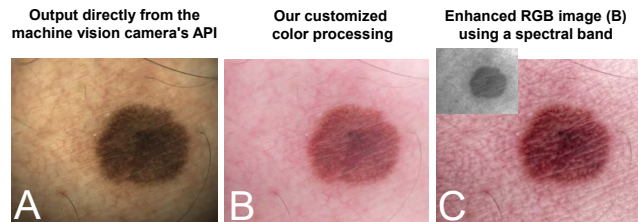

Figure 1. (A) Image directly obtained using the machine vision camera's API. (B) Result from our customized camera processing pipeline after color calibration and photo-finishing. (C) Enhancement of the visible imaging using a selected spectral band to highlight melanin pigmentation. The narrow band image is shown as an inset.

tion beyond simple per-channel gain manipulation. In addition, machine vision cameras lack onboard photo-finishing that allows consumer-oriented cameras to produce perceptually pleasing images. When machine vision cameras' images are used in applications where the image content needs to be displayed to users (especially non-vision experts), the perceived quality of the machine vision imagery often appears lower than consumer-camera imagery – even though the overall sensor performance of the machine vision camera is better (see Figure 1).

In addition, for applications that benefit from images within specific spectral bands – for example NIR spectral bands – there are a range of different options regarding how to use this information to enhance a visible three-channel RGB image. In this paper, we explore these options within the context of dermatological imaging.

Contribution This paper describes how to customize a machine vision camera image pipeline to produce high-quality perceptual images for dermatological applications. As part of this work, we describe the features provided by a typical machine vision camera and discuss why they are not suitable to produce high-quality perceptual images. Our primary contribution is to overview a calibration procedure that not only colorimetrically calibrates the machine vision camera but also allows it to mimic the photo-finishing applied on consumer cameras. Our dermatological imaging

A Schematic drawing of the device

LED light ring
Broadband visible light and
selected non-visible spectral
bands

Machine vision camera

19mm opening

Macro lens

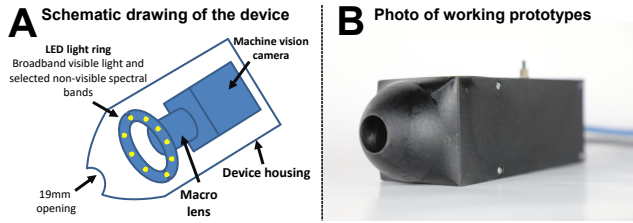Device housing

B Photo of working prototypes

Figure 2. Our device: (A) Schematic diagram of our imaging device. A machine vision camera is integrated inside a closed housing with a ring of LEDs. The LEDs emit a range of visible and non-visible spectra targeting different bio-markers for dermatological disorders. (B) Image of the current prototype device.

system also has the ability to capture images from select narrow spectral bands in both the visible and NIR range. As part of this case study, our secondary contribution is to describe how select spectral images can be incorporated into the visible RGB image to provide enhanced imagery. To this end, we gathered feedback from clinicians to understand their preference regarding which bands and integration methods are preferred.

2. Imaging Device and Application

The customization described in this paper is for a dermatological imaging device that is undergoing clinical evaluation. Figure 2-(A) shows a diagram of the device's construction. A machine vision camera using a Sony IMX172 12MP RGB sensor is integrated into a closed housing with an LED light ring. The light ring is composed for eight (8) LEDs that emitted selected visible and NIR wavelengths, each targeting different biomarkers (e.g., eumelanin and hemoglobin). The corresponding LED wavelengths range between 400-1100 nm. An additional LED that emits broadband visible light is included. The sensor lacks a hot mirror that is typically found on consumer cameras to block NIR light. By triggering image capture with the LED emission, a visible RGB image in addition to the multiple spectral bands can be captured in a single image capture session. Figure 2-(B) shows an image of the actual device.

The visible and non-visible images obtained from this device are used with a proprietary deep-learning module that provides dermatologists with objective data to decide if a biopsy is needed for further investigation of a skin lesion. In addition, the doctors can use the device for magnification of the target and their own visual inspection. When used manually, our device can serve as a potential replacement, or auxiliary device, for existing hand-held dermatoscopes that use a magnifying lens with an illumination source. We note that this paper does not discuss the associated deep-learning system. Instead our focus is on the calibration of the imaging rig, with emphasis on processing the machine vision data for perceptual output.

3. Related Work

Related work is discussed in two areas: (1) camera color calibration and (2) spectral image fusion.

Camera calibration and customization RGB camera sensors have their own spectral sensitivity for each color channel. As a result, images captured by a camera are in a sensor-specific RGB color space. An insightful analysis of the diversity of spectral sensitivities and sensor-specific color spaces for a wide range of cameras can be found in [15]. A minimally processed camera image is typically referred to as a raw-RGB image. Most machine vision cameras and consumer cameras now allow access to the raw-RGB image. A desirable property of raw-RGB images for many computer vision tasks is that their response is linear with respect to scene irradiance.

Colorimetric calibration of a camera sensor is the process of computing the mapping of the raw-RGB color space to a canonical perceptual color space, generally the CIE 1931 XYZ color space or one of its derivatives. Calibration is generally performed using a color rendition chart (e.g., an X-Rite Color Chart) that has color patches that have known CIE XYZ values. By imaging a color chart, simple linear regression can be used to estimate 3×3 matrix transforms based on raw-RGB and CIE XYZ color correspondences. Early work by Hong et al. [13] suggested a color space transform based on higher-order polynomial terms could be used to provide a more accurate mapping. Recent methods, such as Funt and Bastani [3] and Finlayson et al. [8, 9], have developed approaches to estimate the colorimetric transform when the color rendition chart is non-uniformly illuminated. Finlayson et al. [11] introduced polynomial color correction by using a root polynomial that makes the high-order transform invariant to camera exposure. Karaimer and Brown [17] and Bianco et al. [5] proposed various weighting schemes for the color space transform that also consider the scene illumination. These aforementioned approaches all target cameras used in standard imaging scenarios.

There is significantly less work published on customized color calibration of applied imaging rigs for specific tasks. Notable examples include Berman et al. [4], who developed an imaging pipeline for underwater imaging. This work considered the spectral attenuation that occurs in water media. Work by Karaimer and Brown [16] introduced a camera imaging software pipeline that allowed camera emulation starting with a raw-RGB image input. While their work did not target a specific application, the software allows easy customization of the overall camera pipeline for any task. The lack of a concrete example that details how to calibrate a machine vision camera beyond simple color mapping is one impetus for this paper.

Spectral image fusion There are a number of methods suitable to perform image fusion. Early work typically re-

Typical pipeline for machine vision cameras

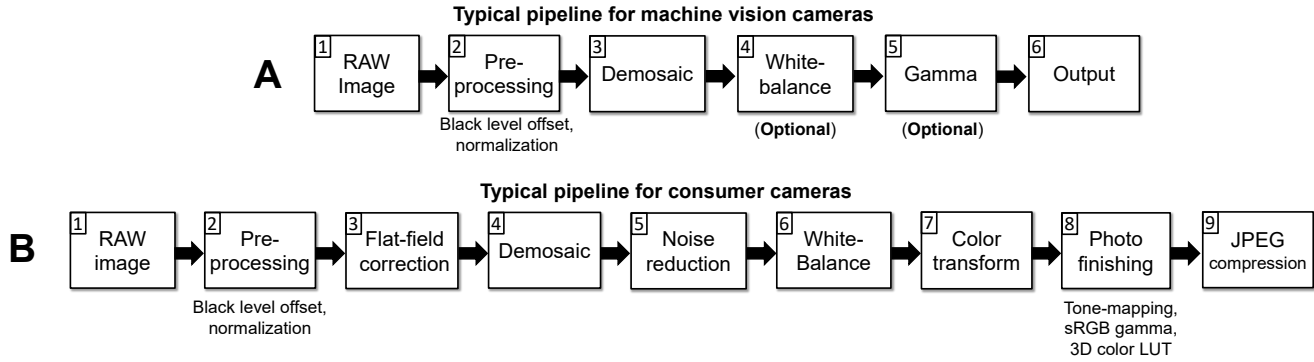**Typical pipeline for consumer cameras**

Figure 3. (A) A typical camera imaging pipeline overviews the common steps applied onboard a machine vision camera. (B) A typical camera imaging pipeline overviews the common steps applied onboard a consumer camera.

lied on copying data from frequency decompositions between images and then reconstructing a new image (e.g., [7, 20]). More modern methods involve joint-image filtering (e.g., [21, 22]). While these methods are generic in nature, they are suitable for transferring narrow band spectral data to RGB images.

There are several works focused directly on processing visible images with the help of non-visible data. For example, Krishnan and Fergus [19] showed how noisy low-light images could be denoised using a corresponding high-quality NIR image. Similar work by Wu et al. [24] removed light glows for low dynamic range web cameras using NIR imagery. Fredembach et al. [12] and Süsstrunk et al. [23] proposed methods that used NIR images to remove skin imperfections like freckles, pores, warts, and wrinkles from RGB images. Zhang et al. [25] showed how to combine NIR images to create high-dynamic-range RGB images. These methods employed variations on joint-image filtering as discussed above. Recent work by Connah et al. [6] and Finlayson et al. [10] proposed gradient domain approaches that avoid both filtering and frequency decomposition.

In this paper, we evaluate several of the methods above to test their suitability for our application for dermatological evaluation.

4. Machine Vision versus Consumer Camera Pipelines

Here we describe the processing steps found on machine vision cameras and consumer cameras. This section helps to reveal why machine vision cameras' images appear visually different from consumer cameras. The section also concludes with a short discussion on common issues that arise when working with machine vision cameras.

Machine vision camera pipeline Figure 3-(A) shows a diagram of a machine vision camera pipeline. Although machine vision cameras are generally more expensive than consumer cameras, the onboard processing is typically much simpler. The following steps are typical of most machine vision cameras.

Step 1: Raw-RGB capture. The raw image is obtained from the camera sensor in a mosaiced Bayer pattern format. The data is typically 10–16 bits per channel.

Step 2: Pre-processing. The raw image is linearized to transform its values to range between 0 and 1. This includes a black-level offset correction based on the camera's current sensor readings.

Step 3: Demosaicing. The image is demosaiced into three full-sized channels.

Step 4: White balance. An optional white balance can be applied. This is typically applied as an independent gain on the red and blue channels only. Most machine vision cameras will provide an auto white balance function or allow this to be manually set via software API calls. Note that most APIs do not call this white-balance.

Step 5: Gamma. Most machine vision cameras will allow an optional gamma function to be applied to the RGB image.

Step 6: Output. The final output image can be optionally quantized to 8-bit, but is typically not compressed.

It is important to note that the machine vision camera image is still in the sensor-specific RGB space. As a result, the image is not suitable for display on monitors which assume a standard RGB (sRGB) display-referred color space. Figure 1-(A) shows an example of a gamma. Although the image may appear correct, the colors are being incorrectly interpreted as sRGB.

Consumer camera pipeline This section provides a brief overview of a consumer camera imaging pipeline; for more details, readers are encouraged to have a look at [16, 1]. Figure 3-(B) shows a diagram of a consumer camera pipeline. Consumer cameras aim for perceptually pleasing images. Steps 1 and 2 are the same as a machine vision camera and are not discussed here. Note that some of the steps may be applied in different orders.

Step 3: A flat-field correction. A spatially varying gain is applied to correct for uneven light fall on the sensor due to the camera's form factor and lens characteristics.

Step 4: Demosaicing. This is applied similarly to the machine vision camera. Often edge sharpen may be included to enhance the appearance of the image.

Step 5: Denoising. Many consumer cameras, especially smartphone cameras, incorporate some type of image denoising.

Step 6: White balance. Unlike machine vision cameras, this is not optional. All consumer cameras apply a white-balance step. White-balance compensates for the scene illumination and prepares the image for a subsequent color space transform to map the sensor-specific raw-RGB to a device-independent perceptual color space.

Step 7: Color space transform. Consumer cameras convert the raw-RGB color space to a perceptual color space based on CIE XYZ, such as a linear standard RGB (sRGB). This transform is reliant on the quality of the previous white-balance step and does have limitations (for more details see recent work by Karaimer and Brown [17]); however, in general, the color transform step serves as a reasonable colorimetric conversion of the raw-RGB image.

Step 8: Photo-finishing. Consumer cameras apply proprietary color manipulation to improve the appearance of the image. While this is shown as a single step in Figure 3-(B), photo-finishing often involves multiple substeps depending on the camera. These steps include per channel tone-curves, application of 3D look up tables (LUT), and selective color manipulation. Spatially varying photo-finishing can also be applied. The photo-finishing step generally imparts a unique look and feel to the final output image that is associated with a particular make and model of a camera (e.g., Nikon, Canon, iPhone, Samsung).

Step 9: Compression and output. Consumer cameras apply JPEG compression to the output image. The image is encoded in a display-referred color space, typically sRGB. These images are suitable for display.

Remarks Many computer vision researchers are not aware of the differences between a machine vision and consumer camera pipeline. Often the API documentation provided with machine vision cameras does not help to clarify these issues. The machine vision camera based on the Sony IMX172 sensor provided additional API calls that allowed manipulation of the captured image in the Hue, Saturation, Value (HSV) space. While HSV can be used to reparameterize an arbitrary RGB color space, it is most often associated with a gamma-encoded sRGB color space. The Sony IMX172 documentation failed to mention that the HSV is in the camera's raw-RGB color space. Only by calling the manufacturer were we able to assess this. In fact, the camera's documentation never specified if a colorimetric conversion was performed onboard the camera.

Not understanding the color space images are saved in can have serious consequences. For example, when displaying sensor-specific RGB images, the observer is not seeing an image that is correctly calibrated for a display. Another serious consequence is when the imagery is used to train machine learning modules. If a deep-learning module is trained using a sensor-specific RGB input, it may not generalize well to other input color spaces – as a result, an application can be restricted to work with only a particular make and model of a camera. The following section discusses how this can be avoided by customizing the camera's pipeline of a device.

5. Custom Camera Pipeline and Results

This section describes our customized camera pipeline for our dermatological imaging device. Further, given the availability of the narrow band spectral data (including non-visible bands), we provide a mechanism to use this information to enhance our RGB image. Figure 4 shows an overview of our customized pipeline.

Flat-field correction Because our LED lights are at different positions, we need to compute a flat-field correction for each LED. We can do this by imaging one of the achromatic patches on a color rendition chart with each LED turned on. This patch has a uniform reflection and therefore any deviation in the captured image is assumed to be due to a combination of light position and lens. To correct the non-uniform response, we construct a map that inversely scales each pixel's intensity to the intensity level in the middle of the image that the lens focuses on best.

Figure 5 shows the vignetting maps computed for each nine LEDs. Due to slight misalignment of the camera's optical axis with the 3D printed housing of our prototype device, the flat-field distortion maps are all dominated by strong fall near the lower region of the imaging field. Such issues should be resolved with proper industrial design of the production-level device.

Color calibration Our sensor needs to be colorimetrically calibrated as discussed in the previous section. We tested both a 3×3 and a 3×11 matrix color space transform (CST) to map our raw-RGB values to the CIE XYZ color space. The 3×11 matrix is based on the polynomial equations proposed by Hong et al. [13]. Since we have a fixed light source in our imaging device, we do not need to compute a white-balance matrix. Instead we can directly compute a transform from the raw-RGB to our target color space. In our case, we map to linear-sRGB. To compute our CST, we place our device over each of the color patches in a Macbeth chart and take a reading with the visible band LED light turned on. Flat-field correction is applied to the captured images before extracting the raw-RGB color values of each patch.

We use iteratively re-weighted least squares [14] to compute the mapping functions based on the color correspon-
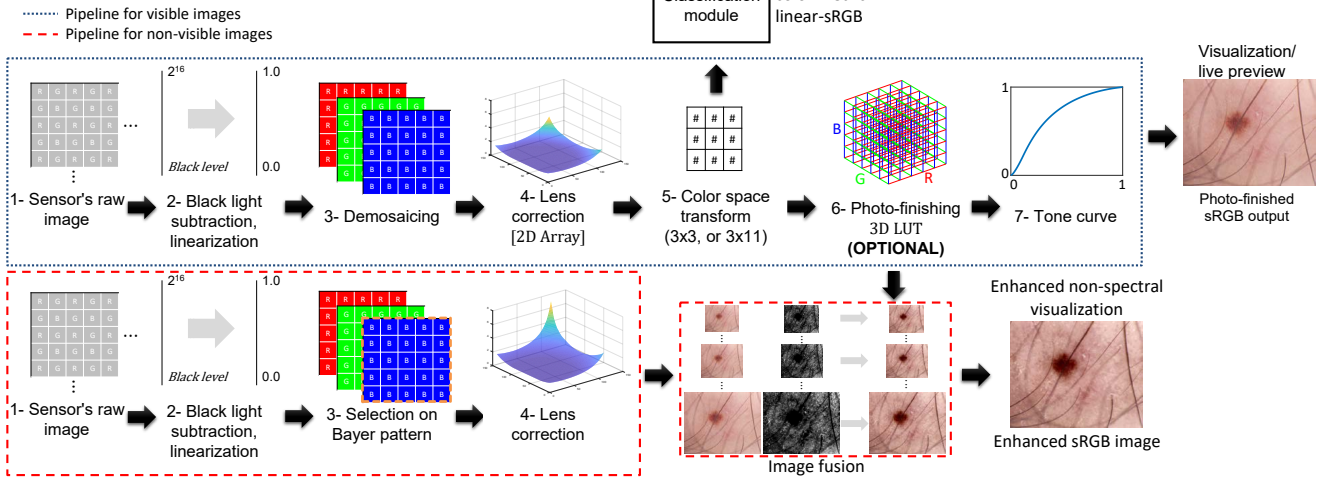
Our customized camera imaging pipeline

Figure 4. This figure shows the customized pipeline design which allows the machine vision camera to behave more similarly to a consumer camera pipeline. A noticeable difference is that we do not need to perform a white-balance step since we will always be imaging under the same lighting conditions. This means the white-balance can be absorbed into the color space transform. We include additional tone curve processing and optional color manipulation (3D LUT) to mimic the appearance of consumer cameras. Our pipeline also includes a select spectral band image pipeline that processes the narrow band spectral image and then integrates it with the RGB image.

dence. To examine the quality of our mapping, we can visualize the angular error between the target patch colors and our mapped colors. Angular error is computed as follows: assuming a target color patch is $I_t = [r_t, g_t, b_t]^T$ and the mapped RGB value is $I_u = [r_u, g_u, b_u]^T$, the angular error E_{ang} (in degrees) between the two colors is:

$$E_{ang} = \frac{180}{\pi} \cos^{-1} \left(\frac{I_{gt} \cdot I_u}{\|I_{gt}\|\|I_u\|} \right), \quad (1)$$

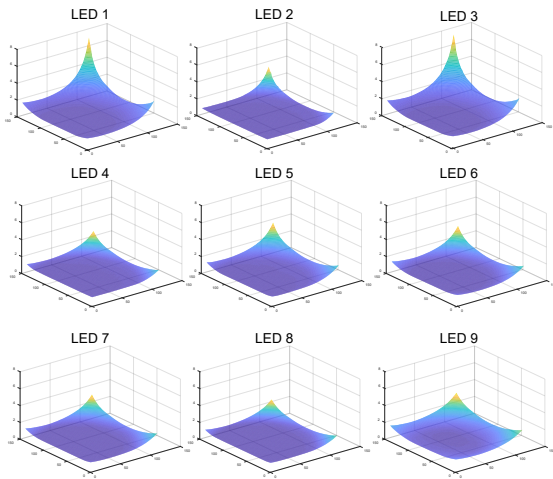where T is the transpose operator and $\cdot$ is the dot product.

Figure 5. Flat field correction maps corresponding to the nine different LED lights.

Angular error is often used to measure color differences to account for potential intensity differences between the two measurements [17].

We also examine our estimated CSTs on a different set of color patches consisting of 128 skin colors obtained from the *Munsell Book of Color*. The *Munsell Book of Color* provides patches that are indicative of skin tones. We also compare the results of applying white balance to the raw-RGB image. The white-balance gains are computed by measuring an achromatic patch with the device and then adjusting the camera's red and blue gain such that $R = G = B$ for all pixel intensities in that patch. The angular RGB errors for each of these methods are shown in Figure 6. We can see that colorimetric calibration drastically improves the accuracy of mapping to a perceptual space.

Photo-finishing After colorimetric conversion, we can send the linear-sRGB images off to our AI module for image analysis. However, for visualization, we still need to apply an additional non-linear mapping. The sRGB standard specifies a 2.2^{-1} gamma encoding to brighten the image. However, all cameras apply a proprietary tone-curve to process the image instead of the 2.2^{-1} gamma encoding. These tone curves can often be extracted from metadata saved in the Adobe DNG file format [16]. Additionally, many cameras, especially DSLRs, include a more complicated color manipulation based on a 3D LUT. The LUTs are associated with different picture styles settings on the camera (see [18]).

For our customization pipeline we are able to do both. Figure 7 shows an example. Figure 7-(A) shows the stan-
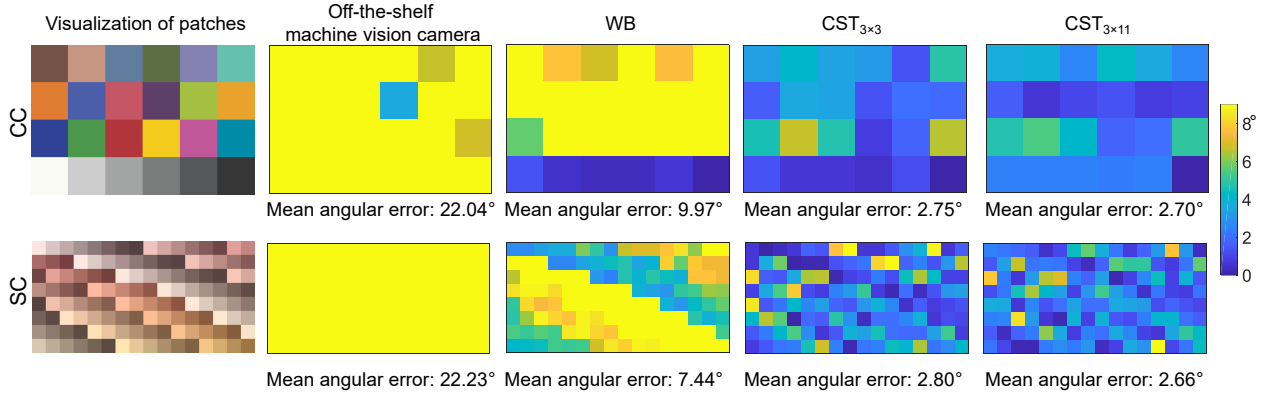
Figure 6. Angular errors for color chart colors (denoted as CC), and on skin colors (denoted as SC) from *Munsell Book of Color*.

dard output of the machine vision camera. Figure 7-(B) shows the linear-sRGB result. Figure 7-(C) shows a tone-curve applied based on Adobe's standard photo-finishing tone-curve. This tone-curve is not associated with any particular camera, but is used by Adobe's Lightroom software when no tone-curve is specified. This produces a perceptual image that looks much more similar to a consumer camera.

Figure 7-(D,E,F) uses meta-data extracted from DNG files to emulate the following camera modes: D-Nikon Vivid mode, E-Canon Portrait mode, and F-Olympus Natural mode. For our working system, we use the Adobe tone-curve (or a custom curve), however; when desired, we can produce outputs that mimic existing camera models. Our colorimetric conversion and photo-finishing can be performed in real time and are part of the live-view display feed from the camera based on the visible RGB image.

6. Spectral data fusion

As part of our customized pipeline, we allow the RGB image to be enhanced using selected spectral bands. Which band to use for enhancement is currently manually specified. Note that this process cannot yet be performed in real time, but instead takes around 1.5–3 seconds per image. This part of our setup is currently not integrated onboard the camera processor, but can be incorporated in future designs.

As shown in Figure 4, our device must first capture and process the single-channel spectral bands using a simplified pipeline. Since these images are not obtained from the visible spectrum, there is no need to perform colorimetric calibration. Moreover, these bands are not demosaiced, but instead the channel most sensitive to the imaged spectra is used (e.g., the raw-B channel for the lights corresponding shorter wavelengths, the raw-R channel for longer wavelengths, such as NIR). Once the image is captured, flat field correction is applied and the images are upsampled to the RGB image's size. Using this image we can perform enhancement. Figure 8 shows the visible and the correspond-

ing spectral images for two sample skin lesions.

We have examined the following three spectral (3) fusion methods. Each is based on a different strategy, including bilateral filtering, Laplacian pyramid filtering, and wavelet-based image fusion.

(1) Bilateral filter [12]: The first method that we use is based on the work by Fredembach et al. [12]. While the authors proposed a method that fused NIR to a visible image to remove skin imperfections, we used the spectral data instead to boost the nonvisible details. This method is based on the fast bilateral filter method introduced by Paris and Durand [21], which is applied on both the visible luminance image and single-channel spectral image. The method works by decomposing the photo-finished sRGB image into the YCrCb color space. A detail layer is obtained from the spectral image by subtracting the input image from its filtered version. This layer is added to the filtered Y channel of the sRGB filtered image. The modified Y channel is recombined with the CrCb channels to produce the final output with enhanced details from the spectral image.

(2) Fast local Laplacian filters [2]: Our second method is based on the fast local Laplacian filters. This method was proposed by Aubry et al. [2], who provide a performance modification over local Laplacian filters first introduced by Paris et al. [22]. While local Laplacian filters were not used for image fusion, the method is fairly straightforward to allow sharing of information between two images. Specifically, we use the algorithm outlined in [2] to compute nine (9) levels of the Gaussian pyramid for the spectral image, as well as the corresponding Gaussian pyramid and the Laplacian pyramid of the visible image's Y channel. Afterwards, we remapped the narrow band image using the S-shaped curve as done in [2] to boost the local gradients of the spectral image. These boosted gradients are blended with the local gradient information from the Y channel of the RGB image. Before we collapse the combined Laplacian pyra-

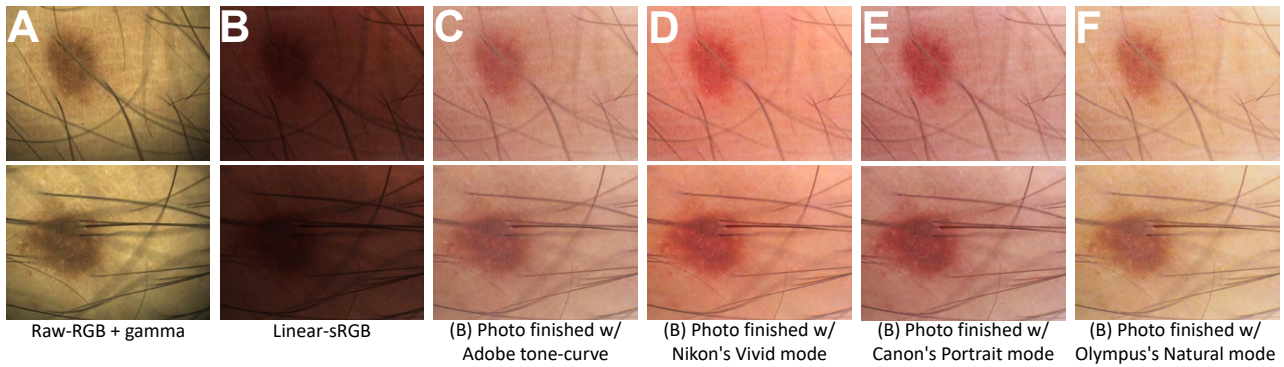| Raw-RGB + gamma | Linear-sRGB | (B) Photo finished w/ Adobe tone-curve | (B) Photo finished w/ Nikon's Vivid mode | (B) Photo finished w/ Canon's Portrait mode | (B) Photo finished w/ Olympus's Natural mode |

Figure 7. This figure shows the output of the machine vision camera with a gamma mapping applied (A). (B) shows the linear-sRGB mapping computed from the colorimetric calibration. (C, D, E, F) show various versions of photo-finishing applied to image (B) to produce different appearances. These photo-finishing versions can be emulated in our custom pipeline.
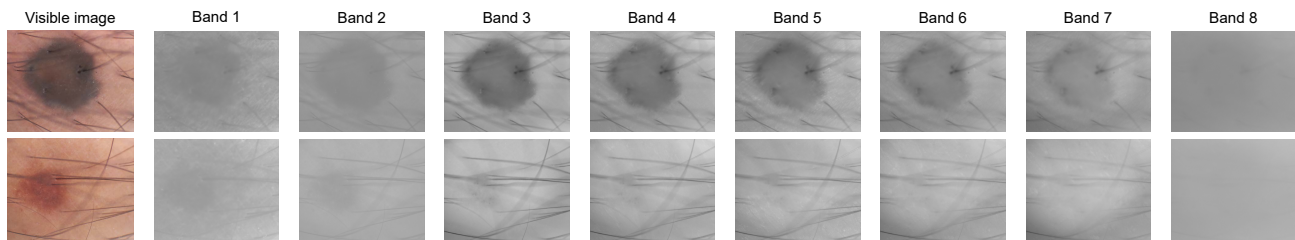

Figure 8. The visible and corresponding narrow band images for two different melanocytic nevus lesions. The first lesion (top row) has a structure that extends deeper into the skin; this is apparent in the individual spectral bands. The second lesion (bottom row) is primarily on the outer layer of the skin and contributes very little to the spectral bands.

mid, the last level of the Gaussian pyramid is copied from the visible Y channel. This effectively transfers the spectral image's details to the visible image Y channel. The modified Y channel is recombined with the CrCb channels to produce the final output.

(3) Wavelet-based image fusion: Our last method is a wavelet-based image fusion approach that was also discussed in [12]. This method also starts by converting the photo-finished sRGB image to the YCbCr color space. A wavelet decomposition is applied using a Symlet wavelet generator on the Y channel of the visible image and the selected spectral image. After the coefficients are calculated, the wavelet coefficients in the visible image's Y channel are replaced by the corresponding spectral image's wavelet coefficient if its coefficient's magnitude is larger. Reconstructing the image with the fused wavelet coefficients performs the image fusion on the Y channel. The new Y image is recombined with the CrCb channels of the visible image to produce the final output.

Figure 9 shows the results of the three approaches. While these results are subjective in nature, our observations are that the local Laplacian filter and bilateral filter methods give similar performance. The wavelet-based methods tend to introduce low frequency intensity shifts in the overall im-

age. The bilateral and local Laplacian filter methods both run in about 2 seconds on a standard workstation PC. The narrow band spectral fusion is still an area of ongoing investigation and will require several rounds of user study to determine the most preferred methods by practitioners. However, regardless of which method will be eventually be adopted, the overall processing pipeline for this image fusion will be based on our current customized pipeline design.

Clinical feedback In order to find out if there was a preference between the methods and the different spectral bands, we conducted a simple study with practicing dermatologists who have been evaluating the device.

The clinicians examined the visible image and the corresponding spectral images as well as the three fusion methods. We showed ten cases and for each case we asked three questions:
(1) Among the three methods used to perform spectral image fusion, do you have a preferred method?
(2) Is there a particular spectral image that you feel provides the most information?
(3) Do you feel this type of fusion is useful for you in a clinical setting (i.e., would it help you make a more informed decision)?
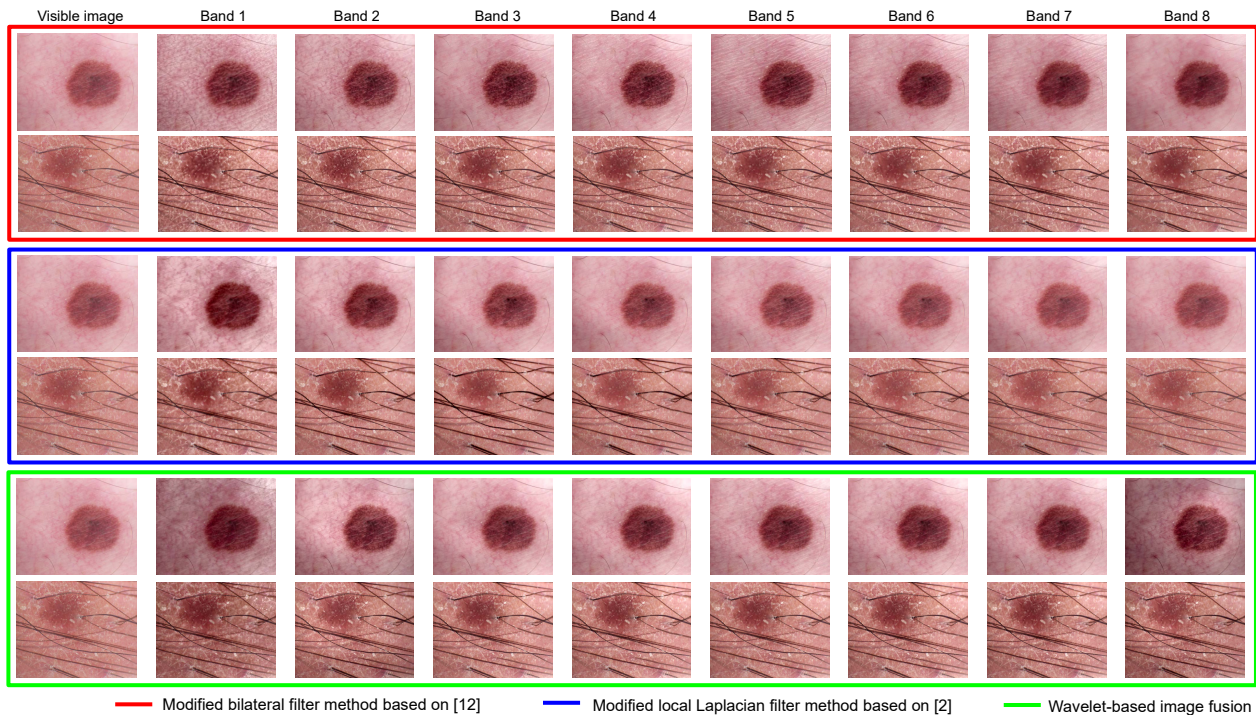
Figure 9. This figure shows the results of several different narrow band spectral images being used to modify the photo-finished sRGB image. The methods applied are the bilateral filter based on a method by Fredembach et al. [12], a local Laplacian filtering approach based on the method proposed by Aubry et al. [2], and a wavelet-based method. While there are subtle differences, the bilateral filter and local Laplacian give similar performance and are preferred over the wavelet-based results.

The feedback was as follows. For question 1, dermatologist 1 chose the wavelet-based method eight times out of ten samples. Similarly, dermatologist 2's preference was the wavelet-based method, which was chosen seven times out of ten examples. Dermatologist 3 chose the method based on bilateral filtering of all their choices.

Regarding which wavelengths to fuse, band 8 (the highest NIR band) was the most preferred. Dermatologist 1 preferred band 1 (the lowest UV band) seven times out of ten cases. Dermatologist 2 equally chose band 2 and band 8 three times out of ten samples, and dermatologist 3's preference was band 8 nine times out of ten examples. This suggests that the NIR data, especially for longer wavelengths, provided useful information to the clinicians.

For the last question, comments from our participants include "Yes, I think that the fusion is very helpful" and "I could see the pattern of each lesion much better (reticular, dots, borders)." One dermatologist who was neutral on the idea of fusion commented "only in certain cases." While the preferences of the dermatologists were not always consistent with each other, their individual preferences were consistent.

7. Concluding Remarks

This paper has detailed a custom in-camera processing pipeline for a dermatological imaging device that captures both visible and selected narrow band spectral information. As part of this work, we have described the basic camera processing pipelines found on machine vision cameras and discussed limitations that are often overlooked in applications of this nature – namely, the lack of colorimetric calibration and photo-finishing on machine vision cameras. We have provided a detailed description of how to calibrate our camera to improve the colorimetric accuracy as well as mimic the photo-finishing routines found on consumer cameras. While our work targets a dermatological device, we believe information described in this paper will be useful for the design of similar devices.

Specific to our dermatological application, we are still exploring ways to integrate the narrow spectral bands for image enhancement. Currently we use only a single spectral band at a time. Combining multiple bands may allow doctors to assess skin lesions faster. Addressing these issues is part of ongoing work. Based on our pilot user study, however, given the diversity in opinions regarding which bands are most useful, we advocate allowing manual selection over an automated strategy.

References

[1] Andrew Adams, Eino-Ville Talvala, Sung Hee Park, David E. Jacobs, Boris Ajdin, Natasha Gelfand, Jennifer Dolson, Daniel Vaquero, Jongmin Baek, Marius Tico, Hendrik P. A. Lensch, Wojciech Matusik, Kari Pulli, Mark Horowitz, and Marc Levoy. The frankencamera: An experimental platform for computational photography. In *SIGGRAPH*, 2010.

[2] Mathieu Aubry, Sylvain Paris, Samuel W. Hasinoff, Jan Kautz, and Frédo Durand. Fast local laplacian filters: Theory and applications. In *SIGGRAPH*, 2014.

[3] Pouya Bastani and Brian Funt. Simplifying irradiance independent color calibration. In *Color and Imaging Conference*, 2014.

[4] Dana Berman, Tali Treibitz, and Shai Avidan. Diving into haze-lines: Color restoration of underwater images. In *BMVC*, 2017.

[5] Simone Bianco, Arcangelo Bruna, Filippo Naccari, and Raimondo Schettini. Color space transformations for digital photography exploiting information about the illuminant estimation process. *Journal of Optical Society America A*, 29(3):374–384, 2012.

[6] David Connah, Mark Samuel Drew, and Graham David Finlayson. Spectral edge image fusion: Theory and applications. In *ECCV*, 2014.

[7] Paul de Zeeuw, Gema Piella, and Henk Heijmans. A matlab toolbox for image fusion (matifus). Technical Report CWI-PNA-E0424, Probability, Networks and Algorithms [PNA], CWI, 2004.

[8] Graham D Finlayson, Maryam Mohammadzadeh Darrodi, and Michal Mackiewicz. The alternating least squares technique for nonuniform intensity color correction. *Color Research & Application*, 40(3):232–242, 2015.

[9] Graham D Finlayson, Han Gong, and Robert B Fisher. Color homography color correction. In *Color and Imaging Conference*, 2016.

[10] Graham D. Finlayson and Alex E. Hayes. Pop image fusion - derivative domain image fusion without reintegration. In *ICCV*, 2015.

[11] Graham D Finlayson, Michal Mackiewicz, and Anya Hurlbert. Color correction using root-polynomial regression. *IEEE Transactions on Image Processing*, 24(5):1460–1470, 2015.

[12] Clement Fredembach, Nathalie Barbuscia, and Sabine Süsstrunk. Combining visible and near-infrared images for realistic skin smoothing. In *Color and Imaging Conference*, 2009.

[13] Guowei Hong, M Ronnier Luo, and Peter A Rhodes. A study of digital camera colorimetric characterisation based on polynomial modelling. *Color Research & Application*, 26(1):76–84, 2001.

[14] Peter J. Huber and Elvezio M. Ronchetti. *Robust Statistics (2nd Edition)*. Wiley, 2009.

[15] Jun Jiang, Dengyu Liu, Jinwei Gu, and Sabine Süsstrunk. What is the space of spectral sensitivity functions for digital color cameras? In *WACV*, 2013.

[16] Hakki C Karaimer and Michael S Brown. A software platform for manipulating the camera imaging pipeline. In *ECCV*, 2016.

[17] Hakki C Karaimer and Michael S Brown. Improving color reproduction accuracy on cameras. In *CVPR*, 2018.

[18] Seon Joo Kim, Hai Ting Lin, Zheng Lu, Sabine Süsstrunk, Stephen Lin, and Michael S Brown. A new in-camera imaging model for color computer vision and its application. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(12):2289–2302, 2012.

[19] Dilip Krishnan and Rob Fergus. Dark flash photography. In *SIGGRAPH*, 2009.

[20] Michel Misiti, Yves Misiti, Georges Oppenheim, and Jean-Michel Poggi. *Wavelets and Their Applications*. Wiley-ISTE, 2007.

[21] Sylvain Paris and Fredo Durand. A fast approximation of the bilateral filter using a signal processing approach. In *ECCV*, 2006.

[22] Sylvain Paris, Samuel W. Hasinoff, and Jan Kautz. Local laplacian filters: Edge-aware image processing with a laplacian pyramid. In *SIGGRAPH*, 2011.

[23] Sabine Süsstrunk, Clement Fredembach, and Daniel Tamburrino. Automatic skin enhancement with visible and near-infrared image fusion. In *ACM MM*, 2010.

[24] Chen Wu, Ramin Samadani, and Prabath Gunawardane. Same frame rate ir to enhance visible video conference lighting. In *ICIP*, 2011.

[25] Xiaopeng Zhang, Terence Sim, and Xiaoping Miao. Enhancing photographs with near infra-red images. In *CVPR*, 2008.