Combining visible and near-infrared images for realistic skin smoothing

Clément Fredembach, Nathalie Barbuscia and Sabine Süsstrunk
School of Computer and Communication Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
{*clement.fredembach, nathalie.barbuscia, sabine.susstrunk*}*@epfl.ch*

Abstract

Skin tone images, portraits in particular, are of tremendous importance in digital photography, but a number of factors, such as pigmentation irregularities (e.g., moles, freckles), irritation, roughness, or wrinkles can reduce their appeal. Moreover, such "defects" are oftentimes enhanced by lighting conditions, e.g., when a flash is used.

Starting with the observations that melanin and hemoglobin, the key components of skin colour, have little absorption in the near-infrared part of the spectrum, and that the depth of light penetration in the epidermis is proportional to the incident light's wavelength, we propose that near-infrared images provide information that can be used to automatically smooth skin tones in a physically realistic manner.

Specifically, we develop a framework that consists of capturing a pair of visible/near-infrared images and separating both of them into base and detail layers (akin to a low/high frequency decomposition) with the fast bilateral filter. We show that a smooth, realistic, output image can be obtained by fusing the base layer of the visible image with the near-infrared detail layer. This method not only outperforms equivalent decomposition in the wavelet domain, but the results also look more realistic than with a simple luminance transfer. Moreover, the proposed method delivers consistently good results across various skin types.

Introduction

Portraiture is one of the most important aspects of photography and large efforts have been undertaken to enhance such images. In fact, while a number of potential artefacts or unattractive details are induced by the capturing process, e.g., red eyes, blur, colour shifts, a number of such unwanted details are "intrinsic" to the photographed person, e.g., wrinkles, freckles, spots. As a result, many techniques are employed by models and photographers alike to mask, or correct, these less appealing features, from image editing software to print airbrushing.

Current research has had success in correcting the deficiencies due to the capturing process, notably in red-eye removal [8] , face detection for accurate focusing [6], and skin tone enhancement via correct white balancing. Local improvement of the skin is, however, still done in a time-consuming manual way, either pre or post image capture (e.g., the application of make-up or the use of specialised software).

The focus of our work is to provide a method to automatically remove or attenuate the visibility of skin features that can be deemed less appealing. The two main constraints of such an approach are that skin tones ought to look realistic after the processing, and that high-frequency details of the images are preserved. Additionally, we do not want to rely on skin detection methods or 3D face modelling, as they can be of limited precision and usefulness when confronted with different skin types and capture conditions [3, 4].

Rather, we propose to look beyond the conventional visible band and into the near-infrared (NIR) part of the electromagnetic spectrum (700-1100 nm). Li et al. [11] showed that NIR portrait images were easier to perform recognition on than visible, colour, images. A comparison of the same portrait in the visible and NIR bands is shown in Fig. 1, where the difference in smoothness can be observed together with the preservation of the face's details. Explanations of these phenomena are given in Section 2.

Figure 1. Visible and NIR representation of a portrait. While the skin appears much smoother in the NIR, all the high-frequency details are effectively preserved.

Our goal, a smooth yet realistic colour image, does, however, require to "fuse" visible and NIR information in an appropriate manner. Fredembach and Süsstrunk [7] proposed that one could obtain colour images with a NIR "look" simply by treating the NIR channel as luminance-related information and substituting it to the original visible image's. While this method performs well on landscape images, the human visual system's sensitivity to colour shifts in skin tones is too high for this method to reliably deliver realistic looking images. Another visible/NIR fusion approach that focuses on the dehazing problem [12] uses a multi-level decomposition of the image to maximise local contrast across the visible and NIR channels. The resulting dehazed images are physically accurate and exhibit a smaller colour shift, but globally the approach is antipodean to our smoothing goal. Indeed, fusion methods generally aim to maximise the amount of information present in the fused image, whereas our goal here is to preserve only relevant details (e.g., eyes, hairs) while removing unwanted ones.

Our approach is as follows. Since NIR images contain all of a face's important features but (almost) none of the unwanted ones, it is a perfect candidate to be a *detail* layer. The visible image, on the other hand, contains the desired brightness and colour information, and is thus an ideal *base* layer. It follows that one can decompose the NIR and visible images into detail and base layers, either in the spatial or in the frequency domain, and then fuse the NIR detail layer with the visible base layer to obtain an enhanced, realistic looking, image. Results show that the fast bilateral filter [5] provides a good decomposition method, and that realistic results are obtained across various skin types.

Skin behaviour across wavelengths

To correctly understand and assess the usefulness of NIR in skin imaging, one needs to examine the wavelength dependency of skin appearance. Skin is very complex to model accurately because of the number of different parameters that have to be taken into account. Indeed, it is not merely a planar reflectance, but truly a translucent, inhomogeneous material. Absorption, reflection, and scattering effects all have to be considered. We present here a simplified model that explains skin behaviour in the NIR. For a more thorough study of the different parameters involved in skin appearance, we refer the reader to [9].

The schematic model of Fig. 2 depicts the structure of skin. The two main areas of interest for imaging are the epidermis and the dermis. The epidermis, also called "melanin layer" is responsible for most of the pigmented coloration of the skin. It is the melanin concentration that gives the skin a colour that can vary from pale to dark brown. The epidermis layer is otherwise fairly thin and no significant scattering occurs. Just underneath the epidermis lies the dermis layer. In this layer, significant scattering occurs as well as hemoglobin absorption (giving the skin its reddish shade).

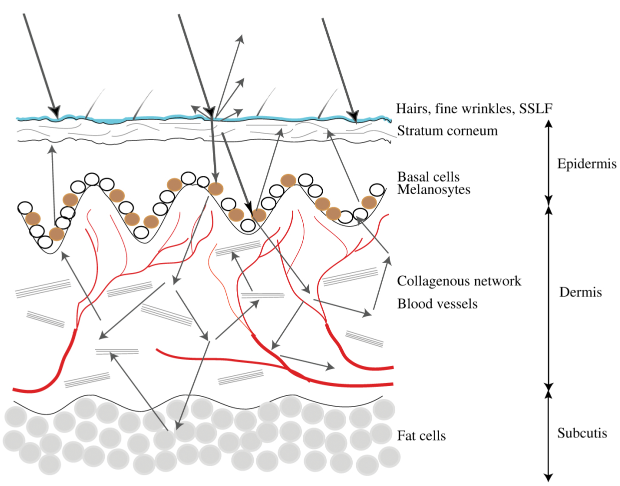

Figure 2. A schematic figure illustrating the different layers of the skin [9].

The main parameters needed to represent light transport in the skin are the absorption coefficient μ_a and the reduced scattering coefficient μ_s, used to express multiple scattering. The absorption coefficients of the epidermis $\mu_{\mathrm{a,epi}}$ and the dermis $\mu_{\mathrm{a,der}}$ can be represented as [10]:

$$\mu_{\mathrm{a,epi}} = f_{\mathrm{mel}}\mu_{\mathrm{a,mel}} + (1 - f_{\mathrm{mel}})\mu_{\mathrm{a,skin}} \quad (1)$$

$$\mu_{\mathrm{a,der}} = f_{\mathrm{blood}}\mu_{\mathrm{a,blood}} + (1 - f_{\mathrm{blood}})\mu_{\mathrm{a,skin}} \quad (2)$$

Where f_{blood} and f_{mel} are the concentration of blood and melanosomes, and $\mu_{\mathrm{a,mel}}$, $\mu_{\mathrm{a,blood}}$, $\mu_{\mathrm{a,skin}}$ are the absorption coefficients of hemoglobin, melanin, and skin layer without any chromophores. All the absorption coefficients are wavelength dependent, $\mu_{\mathrm{a,mel}}$ and $\mu_{\mathrm{a,blood}}$ can be derived from their respective reflectance spectra (see Fig. 3), while the skin layer can be approximated as:

$$\mu_{\mathrm{a,skin}} = 0.244 + 85.3\exp\left(-\frac{(\lambda - 154)}{66.2}\right) \quad (3)$$

Jacques, [10] proposed that the reduced scattering coefficient of the dermis could be approximated as:

$$\mu_s = 2 \times 10^5 \lambda^{-1.5} + 2 \times 10^{12} \lambda^{-4} \quad (4)$$

i.e., the longer the wavelength, the less scattered the light is.

Looking at (1)-(4) and Fig. 3, one realises that NIR light, due to its longer wavelengths, will be less absorbed and less scattered than visible light, therefore penetrating deeper into the skin layers. As such, NIR skin images will contain less "surface features" than normal colour images, a key quality for our smoothing algorithm.

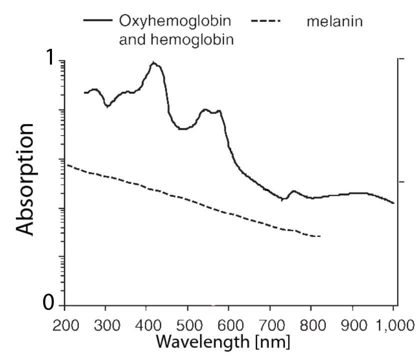

Figure 3. Absorption spectra of hemoglobin and melanin [9].

Near-infrared skin images

In this section, we explore in greater details the different undesired skin "features" that can be smoothed using near-infrared images. Specifically, we look at the reasons why NIR images are ideal candidates, and illustrate the potential improvement that can be obtained by using them in conjunction with traditional colour images.

Pigmentation

Under pigmentation, we refer to the dark brown marks that can be found on the surface of the skin. These details, called freckles, spots, or moles, depending on their thickness, are pigmentary deposits of melanin with well defined contours over the skin layers. Such pigmentation structures generally become more frequent with age, or excessive exposition to UV light, but are also often present in people with fair skin. As such, the contrast between these melanin deposits and the skin can be striking and undesired.

As shown in Fig. 3, melanin does not absorb NIR wavelengths nearly as much as visible light. Consequentially, NIR images of these regions will not display these melanin-rich regions, resulting in a "freckle-free" image. The exact degree of attenuation of pigmentation nonetheless depends on the thickness of the deposits, a feature illustrated in Fig. 4.

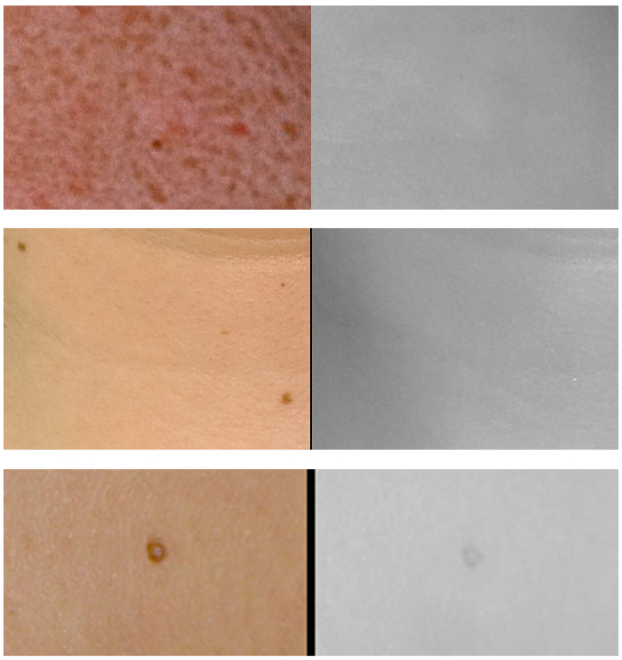

Figure 4. Visible and NIR image pairs of freckles, spots, and moles. Despite the stark contrast provided by these features in the visible image, the NIR representation is essentially melanin-free.

Shallow wrinkles

Wrinkles, distortion of the dermis caused by water-loss or a loss of elasticity, are among the most conspicuous and less-desired features of human skin. The depth of wrinkles vary according to their location and the age of the person. Eye corner wrinkles for instance (perhaps the most common one in portraits) are, on average, 100μm deep for people in their twenties, but almost 700μm deep for those aged 60 and older (although the variance is significant), as measured by Akazaki et al.[1].

To put these numbers in perspective, we look at light propagation in skin. The lesser scattering and absorption at longer wavelengths (see previous section) implies that the depth of penetration of incident light is wavelength-dependent. Specifically, at 550nm (the local maxima absorption of hemoglobin), light penetrates up to 400μ m in the skin. At 800nm, however, this figure is closer to 850μm [2]. As a result, this increased depth will decrease the visibility of shallow wrinkles (since the relative difference of skin depth is less important than the light's intrinsic penetration). This phenomenon is illustrated in Fig. 5 where the perceived strength of wrinkles is noticeably diminished.

Of course, the potential improvement induced by using NIR

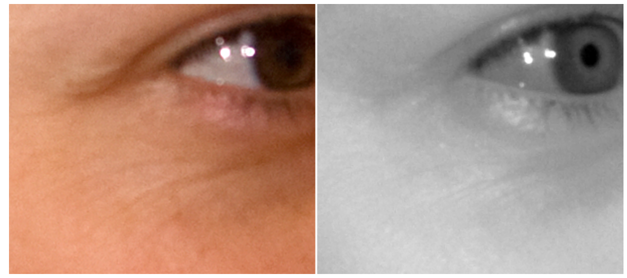

Figure 5. Eye corner wrinkles. The NIR image, due to greater light penetration in skin, appears smoother.

data decreases when the wrinkles are deeper or in the case of superposed skin layers (e.g., eyelids). In this case, some smoothing still occurs, but the features stay mostly unchanged.

Skin texture smoothing

Deeper light penetration combined with the relative transparency of hemoglobin and melanin to NIR allow for more unwanted features to be smoothed out in the NIR image. Flushed skin, visible capillaries, rash, and acne are all features that are visible on the skin surface but whose visibility is decreased in the near-infrared; see the disappearing redness around the eyes in Fig. 6.

Detail preservation

While small scale skin features are generally attenuated in the near-infrared, images that contain skin also contain other details that should be preserved. For instance, in portrait images, high-frequency features such as the distinction between skin, eye, iris, and pupil has to remain. Additionally, hair-based features (e.g., hair, eyelashes, beard) have to remain as sharp as in the original image. A key advantage of NIR images is that they indeed preserve such important details (see Fig. 6). It follows that one can thus use NIR information to obtain a realistically enhanced image.

Enhancing portraits using near-infrared information

A common trait of the undesirable skin features discussed previously is that they all have well-defined contours in the visible, but almost none in the near-infrared image. Colour-wise, freckles and moles share the skin coloration (as it is a melanin concentration issue), and so do other, hemoglobin-related, features. Finally, wrinkles are basically edges within an otherwise low-frequency skin region.

The combination of these characteristics, added to the fact that NIR images preserve otherwise relevant details, make the visible image luminance a prime candidate for NIR modification. In the following, we detail three different methods to incorporate NIR information in the visible image: luminance transfer, wavelet coefficient fusion, and detail layer fusion based on the bilateral filter.

Luminance transfer

In their work on colouring the near-infrared [7], Fredembach and Süsstrunk proposed that the luminance of a regu-

Figure 6. Visible and NIR image pairs of important facial details that have to be preserved. The NIR image contains identical information at a similar level of sharpness.

lar colour image could be replaced by NIR information. This method functions well for most outdoor scenes, albeit with a noticeable shift in colour when the luminance difference between NIR and visible becomes large. Applying this method to skin images is therefore unreliable since the human visual system is particularly sensitive to colour shifts in that area.

Indeed, the method of [7], illustrated in Fig. 7b, yields mixed results. On one hand, the resulting portrait is smooth, but the resulting image is not a realistic rendering rendition of the original.

Wavelet image fusion

As most undesirable features of portraits have a significant high-frequency component, one may turn to the frequency domain in order to obtain a meaningful decomposition of the image into low-frequency components (the skin itself) and high-frequency ones (skin features, as well as eyes, hairs). To do so, we chose to use the wavelet coefficient fusion technique, useful because of its localisation property [14]. In a first step, the visible RGB image is transformed into the YCbCr colourspace. A wavelet decomposition is then performed on both the Y channel and the NIR image. Coefficients are calculated and, for the high-pass versions of the image, the minimal coefficient between the Y and NIR are kept, and the image is then reconstructed. Specifically, we use here the Haar wavelet generator and four levels of decomposition.

The difficulty in using this method stems from the large variety of the wavelet parameters (number of decompositions, wavelet seed used), as well as the diversity of skin features one aims to identify. As such, this method can be unreliable, as exemplified in Fig 7c, where half of the face is correctly smoothed, but the other one is not, due to a difference in image sharpness

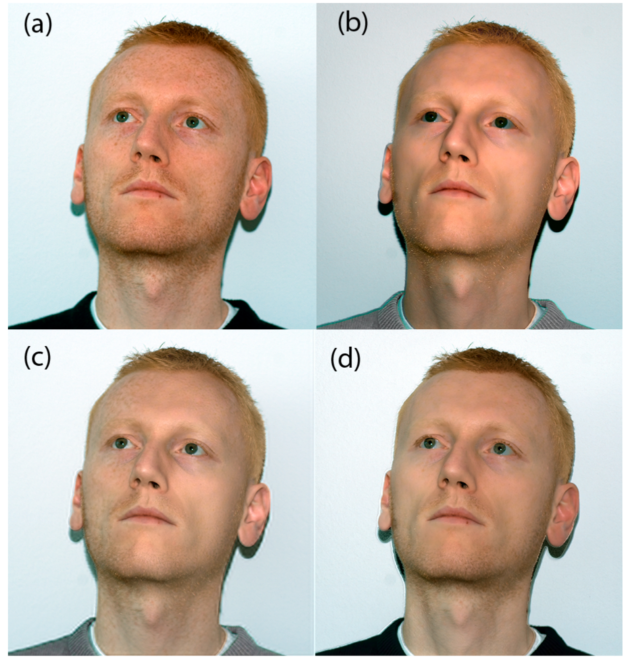

Figure 7. **(a)** The original colour portrait. **(b)** Smoothing with luminance transfer **(c)** Smoothing using wavelet coefficient fusion **(d)** Smoothing using the fast bilateral filter method

(implying that the relevant coefficients will be located in a different sub-band).

Bilateral filter

The bilateral filter, proposed by Tomasi et al. [13] is an edge-aware spatial filter that decomposes an image into base and detail layers. The base layer comprises mostly low frequency information with a small edge content, while the detail layer is primarily high-frequency information. Related to portrait images, the base layer is responsible for large, smooth, regions and the detail one for the stark, small-scale, changes across these large regions (see Fig. 8).

Figure 8. The base and detail layers of the visible portrait of Fig. 7a. Most of the freckles are indeed locate in the detail layer.

Given their size, most, if not all, of undesirable skin features will be located in the detail layer of the visible luminance image. Decomposing the NIR image will, however, yield a detail layer that contains all high-frequency information except for these conspicuously absent undesirable features. A simple method to smooth the image is therefore to fuse the detail layer of the NIR image with the base layer of the visible luminance. Chrominance information is then added, and the image is transformed back into RGB. The entire procedure is illustrated in Fig. 9, and a sample result, shown in Fig. 7d, exhibit a natural-looking smooth image.

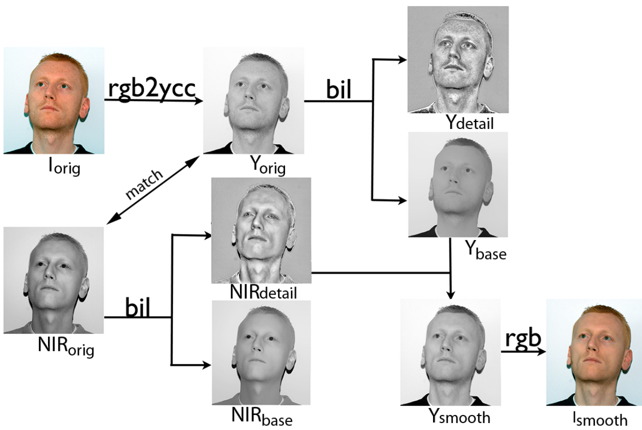

Figure 9. Flowchart detailing the bilateral filter layer merging procedure.

A major drawback of the complete bilateral filter decomposition is its speed, we thus use here the fast bilateral filter proposed by Durand and Dorsey in [5], with no significant decrease in image quality.

Experiments

In this section, we detail the experimental protocol that was used in obtaining the images, as well as the preprocessing steps undertaken. The smoothing is done according to the fast bilateral filter method described in the previous section.

Image capture

All the images of this experiments were captured identically. A modified Canon 400D with a L-series 85mm prime lens was alternatively fitted with a B& W 486 UV-NIR cut filter (for the visible image) and a Wratten 97c equivalent filter (for the NIR image). A flash (speedlight 580 EX) set to 1/4 of maximum power was the sole light source and the camera settings were not modified between the visible and NIR image capture.

The white balancing of the camera was set to manual and designed to capture faithful visible images. As a result, the NIR images exhibited a slight purple colour cast, but that cast was ultimately irrelevant as only the red channel of the raw NIR image was used. The only adjustment between the shots was the focus, in order to obtain images of comparable sharpness. The raw outcome of this procedure is the input of our algorithm.

Even though the subjects were asked to look at a fixed point, and all care was taken to minimise movement, some misalignment of images occurred nonetheless. In this work, we have registered the images using Photoshop's "align layers" function, although fully automatic methods can also be employed.

Brightness adjustment

In order to have comparable detail layers, the NIR image is adjusted with respect to the visible luminance. Specifically, the mean value of the NIR image is shifted towards the one of the visible, and the histogram is then stretched by giving the NIR image a black and a white point. The result can be seen in Fig. 10.

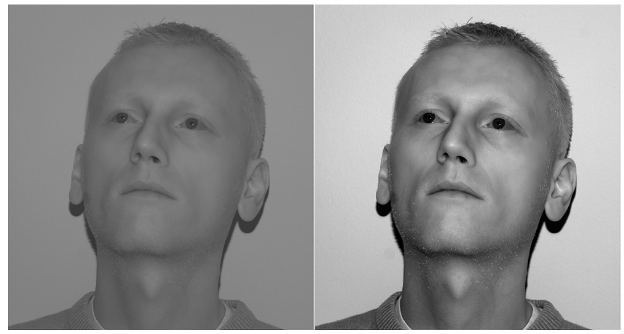

Figure 10. The result of adjusting NIR brightness.

Results

The results presented in Figs. 7d and 11 are all obtained with the bilateral method explained above. Every step is automatic, except the refocusing of the lens at the time of image capture. Importantly, the quality of the results does not depend on the type of skin considered.

Conclusions

Near-infrared wavelengths are not as much scattered or absorbed by skin as visible ones. NIR skin images therefore exhibit much less unwanted "features" than their visible counterparts. Using a standard image capture framework, we propose that the visible luma image and the NIR one can be decomposed into base and detail layer, using the fast bilateral filter. Fusing the detail layer of the NIR image with the base layer of the visible one yields images that are noticeably smoother, yet that preserve important high-frequency details. The proposed algorithm is able to process full resolution images in a matter of seconds (using Matlab) and is reliable, irrespectively of the considered skin types.

Acknowledgement

The work presented in this paper was supported in part by the Swiss National Science Foundation under grant number 200021-124796/1.

References

[1] S Akazaki, H Nakagawa, H Kazama, O Osanai, M Kawai, Y Takema, and G Imokawa. Age-related changes in skin assessed by a novel three-dimentional morphometric analysis. *British Journal of Dermatology*, 147:689–695, 2002.

[2] R R Anderson and J A Parrish. *Optical Properties of Human Skin*, chapter 6. Plenum Pres., 1982.

[3] T Arici and S Dikbas. Skin-aware local contrast enhancement. In *Proc. of the IEEE International Conference on Image Processing*, pages 521–524, 2007.

[4] C Dubout, M Tsukada, R Ishiyama, C Funayama, and S Süsstrunk. Face image enhancement using 3d and spec-

Figure 11. Realistically smoothed skin while preserving important features (the darkish blurred spot present in all the images results from dust being present on the sensor).

tral information. In *IEEE International Conference on Image Processing*, 2009.

[5] F Durand and J Dorsey. Fast bilateral filtering for the display of high-dynamic-range images. *Proc. ACM SIGGRAPH*, pages 257–266, 2002.

[6] H Endo. Imaging device performing focus adjustment based on human face information. US Patent Application 2007/0280662, 2007.

[7] C Fredembach and S Süsstrunk. Colouring the near-infrared. In *Proc. of the 16th IS&T/SID Color Imaging Conference*, Portland, OR, 2008.

[8] M Gaubatz and R Ulichney. Automatic red-eye detection and correction. In *Proc. of the IEEE International Conference on Image Processing*, pages 804–807. IEEE, 2002.

[9] T Igarashi, K Nishino, and S K Nayar. The appearance of human skin: A survey. *Foundations and Trends in Computer Graphics and Vision*, 3:1–95, 2007.

[10] S L Jacques. Skin optics. Oregon Medical Laser Center Monthly news and articles on Biomedical Optics and Medical Lasers. http://omlc.ogi.edu/news/jan98/skinoptics.html, 1998.

[11] S Z Li, R Chu, S Liao, and L Zhang. Illumination invariant face recognition using near-infrared images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29:627–639, 2007.

[12] L. Schaul, C Fredembach, and S. Süsstrunk. Color image dehazing using the near-infrared. In *IEEE International Conference on Image Processing*, 2009.

[13] C Tomasi and R Manduchi. Bilateral filtering for gray and color images. In *IEEE Sixth International Conference on Computer Vision*, pages 839–846, Bombay, India, 1998. IEEE.

[14] P M De Zeeuw, G Piella, and H J A M Heijmans. A matlab toolbox for image fusion (matifus). CWI report: PNA-E0424, 2004.