Tone Mapping for Shortwave Infrared Face Images

Maya Harel and Yair Moshe
Signal and Image Processing Laboratory (SIPL)
Department of Electrical Engineering, Technion – Israel Institute of Technology
Technion City, Haifa, 32000, Israel
yair@ee.technion.ac.il

***Abstract*—Sensing in the Shortwave Infrared (SWIR) range has only recently been made practical. The SWIR band is not visible to the human eye but shows shadows and contrast in its imagery. Moreover, SWIR sensors are highly tolerant to challenging atmospheric conditions such as fog and smoke. However, fundamental differences exist in the appearance between images sensed in visible and SWIR bands. In particular, human faces in SWIR images do not match human intuition and make it difficult to recognize familiar faces by looking at such images. In this paper, we deal with a novel tone mapping application for SWIR face images. We propose a technique to map the tones of a human face acquired in the SWIR band to make it more similar to its appearance in the visible band. The proposed technique is easy to implement and produces natural looking face images.**

***Index Terms*—Shortwave infrared (SWIR), infrared imaging, tone mapping.**

I. Introduction

Historically, the SWIR band (0.9-1.7μm) has been relatively inaccessible for imaging applications due to the lack of large format high sensitivity detectors that respond to those wavelengths. However, recent advances in detector technology have made SWIR imaging practical [1]. The interest in the SWIR band is driven by its advantages relative to other imaging bands, such as the visible band and near infrared (NIR) band. Due to its reflective nature, target signatures in the SWIR band are dominated by reflection of external sources of illumination, much like visible light and opposed to thermal emission of radiation which occurs at the longer infrared wavelengths. Detection of hidden targets is another benefit of SWIR since many man-made materials that have a very different reflectance in the visible/NIR bands have a nearly identical reflectance in the SWIR band and that reflectance is typically very different from the reflectance of naturally occurring background materials. Due to its longer wavelength, SWIR has better penetration through atmospheric obscurants. Therefore, SWIR imaging produces high SNR images in the presence of smoke, mist, fog, etc., as well as under low-light conditions or at night-time. SWIR illumination is invisible to the human eye and is undetectable by silicon-based cameras. On the contrary, a NIR illumination source can be localized by detecting its purple glow, which is observable by a naked eye [2]. All these properties make the SWIR modality suitable for a wide variety of applications, especially when visible spectral images are not feasible.

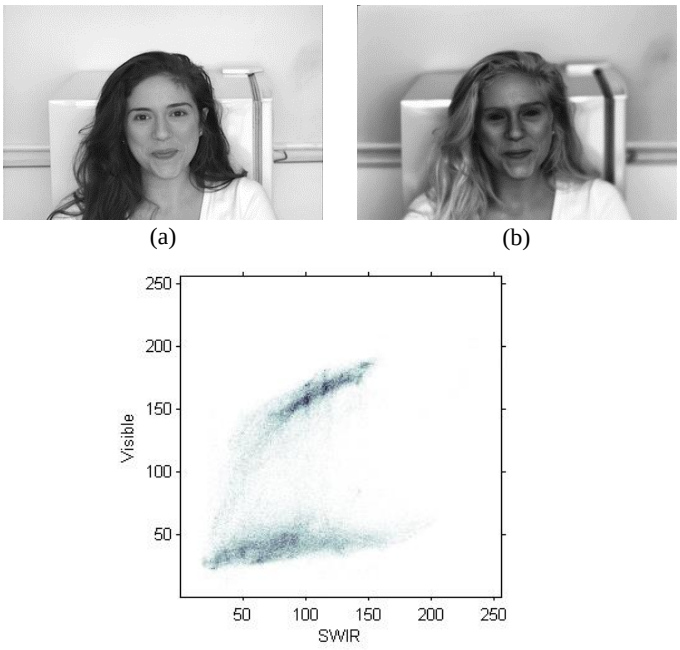
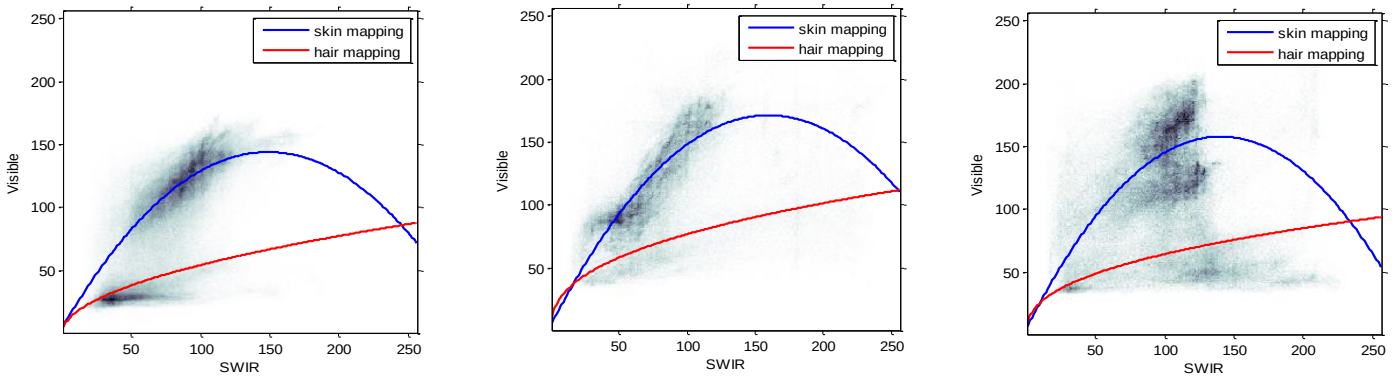


Fig. 1. A subject face acquired in (a) visible band, and (b) SWIR band. Note the fundamental difference in appearance between the two bands. A graph of visible luminance values vs. SWIR values for all face pixels in (a) and (b) is presented in (c). The top cluster represents skin pixels while the bottom cluster represents hair pixels.

Like in the visible band, the appearance of a particular object in the SWIR band is determined by its reflectance and the ambient illumination. However, fundamental differences exist in appearance between images sensed in the visible and the SWIR band. In particular, the appearance of human faces in SWIR imagery differs significantly from their appearance in visible imagery. Due to the strong absorption by water in the SWIR band, the presence of moisture in the target surface has a significant impact on appearance in SWIR images. High moisture content leads to increased absorption of SWIR radiation, which is responsible for dark appearance of surfaces such as human skin that has significant water content [1]. Fig. 1(a) and Fig. 1(b) present two subject faces acquired in the visible and SWIR bands respectively. In the SWIR band, clothing typically takes on a uniform bright appearance, skin is dark and hair includes bright stripes. For many applications,



(a) (b) (c)

Fig. 2. Visible luminance values vs. SWIR values for all face pixels in the dataset acquired (a) indoor with fluorescent illumination, (b) outdoor at daytime, and (c) outdoor at nighttime with street light illumination. The top cluster represents skin pixels and is modeled by a quadratic function. The bottom cluster represents hair pixels and is modeled by a gamma correction function.

recognition of humans is a critical requirement. The ability to do so, both in the visible and the SWIR band, is driven by the ambient illumination available for a particular application as well as the reflectance of face parts. Even for a human viewer, it is difficult to recognize a face in the SWIR band based on its appearance in the visible band.

The utilization of the SWIR sub-band has yet to be studied in depth [3]. Only few previous works in the literature consider the difference in appearance between visible and SWIR images. Some of these works deal with multi-sensor image fusion [4, 5]. Other works deal with extraction of band-invariant features from face images for the task of face verification, detection or recognition across bands [2, 6, 7]. These works use gradient-based or texture-based features and do not try to map the tones of a SWIR image to the tones of its counterpart visible image. Some techniques for mapping image tones were suggested in the area of image colorization. The most relevant works deal with infrared image colorization [8-10]. However, due to the unique nature of SWIR band and due to the texture-based nature of such methods, they cannot be applied here.

In this paper, we show that there exists a tone mapping between visible and SWIR face images and propose a novel tone mapping technique for such images. The aim of this technique is to map the tones of a human face acquired in the SWIR band to make it more similar to its appearance in the visible band.

II. Tone Mapping for SWIR Images

Fig. 1(c) presents a graph of visible luminance values vs. SWIR values for the face pixels in Fig. 1(a) and Fig. 1(b). In this graph, pixel values are arranged in two clusters. The top cluster represents skin pixels, while the bottom cluster represents hair pixels. This arrangement hints that a tone mapping between visible luminance and SWIR value exists and is different for hair pixels and for skin pixels. In order to validate this assumption, we have built a dataset of a few dozen face images. Each subject face was acquired simultaneously by a camera in the visible band and by a Goodrich SWIR camera [11]. Images were acquired indoor with fluorescent illumination, outdoor at daytime, and outdoor at nighttime with street light illumination. Homography matrices between the visible and SWIR cameras were computed based on image key points. The homography matrices were used to compensate for the different viewpoints and camera parameters and in order to align spatially a visible image and its counterpart SWIR image.

Fig. 2 presents three graphs of visible luminance values vs. SWIR values for pixels of all faces in the dataset in the three illumination conditions. Like in Fig. 1(c), pixel values are arranged in two clusters – the top cluster represents skin pixels and the bottom cluster represents hair pixels. Note that, according to the amount of illumination, outdoor at nighttime images contain a large amount of noise, indoor images contain less noise and the outdoor at daytime images are the most noise-free. Other sources of noise are spatial misalignments in compensating for the different viewpoints and camera parameters.

After we have shown that a mapping between visible luminance and SWIR values exists, we will now describe a technique to perform such a mapping. Since skin pixels and hair pixels are typically arranged in two different clusters, we handle them separately. First, we segment the head from its background. Then, we segment the head into skin and hair (including eyebrows) segments. This is currently done manually but can be easily extended to automatic segmentation using techniques such as those suggested in [12, 13].

For skin pixels, we apply a quadratic function:

$$x_{swir'} = a_2 {x_{swir}}^2 + a_1 x_{swir} + a_0, \qquad (1)$$

where a_0, a_1, and a_2 are constants. For hair pixels, expansion of the dynamic range of the SWIR values is required for a natural look. Expansion of the dynamic range is achieved by applying a γ correction [14]:

$$x_{swir'} = b {x_{swir}}^{\gamma}, \qquad (2)$$

where b and γ are constants and $\gamma < 1$. This function has greater sensitivity to relative differences between darker tones than between lighter ones, thus it compensates for properties of human vision.

(a) (b) (c)

Fig. 3. Subject faces acquired indoor with fluorescent illumination in (a) visible band, (b) SWIR band, and (c) SWIR band after tone mapping using the proposed technique.

III. Experimental results

We tested the proposed mapping technique with our face dataset. Optimal mapping parameter values were found by least-squares estimation. This procedure minimizes the sum of squared residuals, where a residual is the difference between visible pixel values and their counterpart SWIR values mapped by the mapping function. Different parameter values were found for each illumination condition. The three resulting mapping functions are presented graphically in Fig. 2. In order to find the optimal parameter values, we used leave-one-out cross validation. For n face images, mapping parameter values were found for $n-1$ images and testing was performed for the n^{th} face image. This procedure was repeated n times, each time for a different selection of a test image.

Fig. 3, Fig. 4 and Fig. 5 present mapping results for indoor fluorescent illumination, outdoor at daytime, and outdoor at nighttime with street light illumination, respectively. For some faces, mapping results have a very natural appearance. For others, mapping results have a slightly unnatural tone or contain some minor errors due to segmentation or spatial alignment inaccuracies. However, for all faces in the dataset, the resulting image has a more natural appearance than its counterpart input SWIR image. The suggested mapping technique enhances facial features and hair tones, and produces an image that resembles an image in the visible band. This allows a human observer to recognize a familiar person by looking at such an image.

In order to quantify our results, we have computed the following measure:

$$NRMSE = \frac{RMSE(x_{swir'}, x_{visible})}{RMSE(x_{swir}, x_{visible})}, \qquad (3)$$

where $RMSE(a,b)$ is the root mean square error between a and b, $x_{visible}$ is a vector of face pixels in the visible band, x_{SWIR} is a vector of face pixels in the SWIR band, and $x_{SWIR'}$ is a vector of face pixels in the SWIR band after tone mapping using the proposed technique. For indoor illumination we got a mean $NRMSE = 0.54$ with variance = 0.07, for outdoor at daytime we got a mean $NRMSE = 0.45$ with variance = 0.05, and for outdoor at nighttime we got a mean $NRMSE = 0.62$ with variance = 0.10. As expected, for all three illumination conditions, the resulting mapped pixel values are closer to visible pixel values than the input SWIR values thus $NRMSE < 1$. When illumination is weak, images are noisy,

(a) (b) (c)

Fig. 4. Subject faces acquired outdoor at daytime in (a) visible band, (b) SWIR band, and (c) SWIR band after tone mapping using the proposed technique.

thus the mean $NRMSE$ is small and its variance among different face images is large.

IV. Conclusion

SWIR cameras have many advantages and an important disadvantage - human faces in SWIR images do not match human intuition and make it difficult to recognize familiar faces by looking at such images. In this paper, we propose a technique for tone mapping SWIR face images in order to make them more similar to their appearance in the visible band. The proposed technique applies two different mapping functions – one mapping function is applied to skin pixels and a different mapping function is applied to hair pixels. The technique was tested with a dataset of subject faces built for this purpose. It is easy to implement and produces natural looking face images that resemble their appearance in the visible band.

Acknowledgment

The authors would like to thank Prof. David Malah, head of SIPL, and Nimrod Peleg, chief engineer of SIPL, for their support and helpful comments. The authors would also like to thank all those people who agreed to include their face images in the dataset.

References

[1] T. Haran, "Short-wave Infrared Diffuse Reflectance of Textile Materials," Department of Physics and Astrinomy, Georgia State Univeristy, 2008.

[2] F. Nicolo and N. A. Schmid, "Long Range Cross-spectral Face Recognition: Matching SWIR against Visible Light Images," *Information Forensics and Security, IEEE Transactions on,* vol. 7, pp. 1717-1726, 2012.

[3] R. Shoja Ghiass, O. Arandjelović, A. Bendada, and X. Maldague, "Infrared Face Recognition: A Comprehensive Review of Methodologies and Databases," *Pattern Recognition*, vol. 47, pp. 2807-2824, 2014.

(a) (b) (c)

Fig. 5. Subject faces acquired outdoor at nighttime with street light illumination in (a) visible band, (b) SWIR band, and (c) SWIR band after tone mapping using the proposed technique.

[4] Z.-u. Rahman, D. J. Jobson, G. A. Woodell, and G. D. Hines, "Multisensor Fusion and Enhancement using the Retinex Image Enhancement Algorithm," in *AeroSense 2002*, 2002, pp. 36-44.

[5] K. Lee, J. Kriesel, and N. Gat, "Night Vision Camera Fusion with Natural Colors Using a Spectral/Texture Based Material Identification Algorithm," in *Military Sensing Symposia (MSS) Specialty Group on Passive Sensors*, 2010.

[6] T. Bourlai, N. Kalka, A. Ross, B. Cukic, and L. Hornak, "Cross-spectral Face Verification in the Short Wave Infrared (SWIR) Band," in *Pattern Recognition (ICPR), 2010 20th International Conference on*, 2010, pp. 1343-1347.

[7] Z. Zhang, D. Yi, Z. Lei, and S. Z. Li, "Regularized Transfer Boosting for Face Detection across Spectrum," *Signal Processing Letters, IEEE*, vol. 19, pp. 131-134, 2012.

[8] A. Toet, "Colorizing Single Band Intensified Nightvision Images," *Displays*, vol. 26, pp. 15-21, 2005.

[9] T. Hamam, Y. Dordek, and D. Cohen, "Single-band Infrared Texture-based Image Colorization," in *Electrical & Electronics Engineers in Israel (IEEEI), 2012 IEEE 27th Convention of*, 2012, pp. 1-5.

[10] M. A. Hogervorst and A. Toet, "Fast Natural Color Mapping for Night-time Imagery," *Information Fusion*, vol. 11, pp. 69-77, 2010.

[11] "Sensor Unlimited Inc., http://www.sensorsinc.com/," 2014.

[12] P. Viola and M. J. Jones, "Robust Real-time Face Detection," *International Journal of Computer Vision*, vol. 57, pp. 137-154, 2004.

[13] C. Scheffler and J.-M. Odobez, "Joint Adaptive Colour Modelling and Skin, Hair and Clothing Segmentation using Coherent Probabilistic Index Maps," in *British Machine Vision Association-British Machine Vision Conference*, 2011.

[14] K. N. Plataniotis and A. N. Venetsanopoulos, *Color Image Processing and Applications*: Springer, 2000.