MVA2011 IAPR Conference on Machine Vision Applications, June 13-15, 2011, Nara, JAPAN

Human Skin Detection by Visible and Near-Infrared Imaging

Yusuke Kanzawa
Toyota Central R&D Labs., Inc.
41-1, Yokomichi, Nagakute, Nagakute-cho
Aichi-gun, Aichi-ken, 480-1192, Japan
kanzawa@mosk.tytlabs.co.jp

Yoshikatsu Kimura
Toyota Central R&D Labs., Inc.
41-1, Yokomichi, Nagakute, Nagakute-cho
Aichi-gun, Aichi-ken, 480-1192, Japan
YKimura@mosk.tytlabs.co.jp

Takashi Naito
Toyota Central R&D Labs., Inc.
41-1, Yokomichi, Nagakute, Nagakute-cho
Aichi-gun, Aichi-ken, 480-1192, Japan
naito@mosk.tytlabs.co.jp

Abstract

Advanced driver assistance systems require fast and robust algorithms for detecting lanes, vehicles, pedestrians, and other objects. However, the conventional methods for detecting objects, which employ pattern recognition, have many problems, for example, an increase in processing time, a difficulty of detecting occluded objects, a necessity of huge data for machine learning and so on. To overcome these problems, we propose a method for detecting human skin that employs a spectroscopic camera. Each substance has its own reflection characteristics. In particular, human skin reflects flesh-colored light in the visible region and absorbs light with a wavelength of 970 nm. Thus, to detect human skin, we developed a six-band camera that has red, green, blue, 870, 970, and 1050 nm filters. By using this six-band camera, we can detect human skin in YC_rC_b color space around the visible region and dips around 970 nm in the near-infrared region. We confirm availability of the proposed method for human skin detection and demonstrate the effectiveness for detecting occluded pedestrians from images in road environment.

1 Introduction

Advanced driver assistance systems (ADAS) [1] have recently been developed to promote safety and comfort. ADAS employ radar sensors and cameras to detect lanes, vehicles, pedestrians, and other objects on the road [2][3][4][5]. Cameras have superior spatial resolutions and can be used to identify specific objects using image processing. Consequently, cameras are used in many ADAS.

ADAS with cameras apply pattern recognition to detect pedestrians and vehicles. They use monochrome or color cameras. However, such systems have many problems, for example, an increase in processing time, a difficulty of detecting occluded objects, a necessity of huge data for machine learning and so on. In an effort to overcome these problems, we are developing object detection methods based on spectroscopy. Each substance has its own unique reflection characteristics [6]. Our goal is to develop a robust object detection method based on spectroscopy for road environment.

In this paper, we first analyze the reflectance characteristics of human skin. Based on this analysis, we next develop a camera that can acquire visible (Vis) and near-infrared (NIR) images. Finally, we propose a robust method for detecting human skin that uses the spectroscopic characteristics of skin.

The remainder of this paper is organized as follows. Section 2 describes the reflectance spectrum of skin. In section 3, we explain our proposed camera developed for acquiring Vis and NIR simultaneously without displacement between the images. Lastly, we describe an algorithm for detecting human skin from Vis and NIR images and we present experimental results for pedestrian detection.

2 Reflectance Spectrum of Human Skin

Spectroscopy is the study of how substances absorb, transmit, or reflect light. We focus on UV–Vis–NIR spectroscopy because a conventional camera sensor with a CCD can obtain images in the ultraviolet (UV), Vis, and NIR regions. UV–Vis–NIR spectroscopy is applied for optical absorbance and reflectance measurements in the wavelength range 200–1500 nm. Because a substance has its own unique reflectance characteristics, we can distinguish a substance by analyzing the light reflectance.

Reflectance spectra of some substances, for example asphalt, road markings and human skin are depicted in Figure 1. The horizontal and vertical axis show the wavelength and reflectance, respectively.

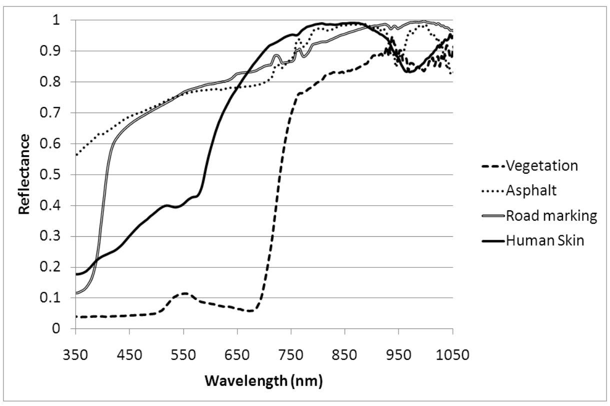

Figure 1. Reflectance spectra of various substances.

Figure 2 shows only a reflectance spectrum of human skin. Skin has a lower reflectance at shorter wavelengths (about 350 nm) than at longer wavelengths (about 1050 nm). In particular, skin has a unique property that it absorbs light around a wavelength of 970nm in the NIR region. We measured the reflectance spectra of skin from about 50 people who have various races; all the spectra revealed that light was absorbed at a wavelength of 970 nm, as shown in Figure 3.

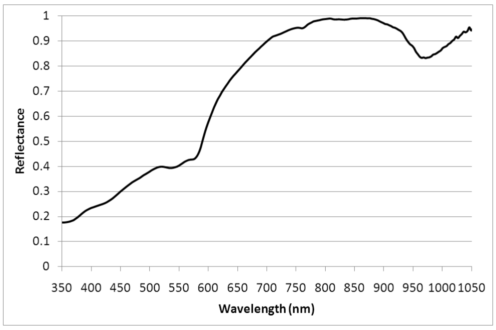

Figure 2. Reflectance spectra of human skin.

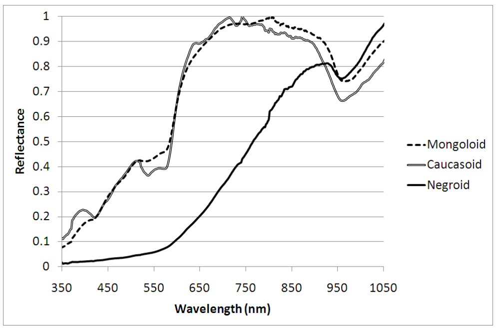

Figure 3. Spectra of various human skin.

3 Six-Band Camera

Human skin reflects flesh-colored light in the Vis region and absorbs light with a wavelength of 970 nm in the NIR region. By exploiting these characteristics, we developed a camera available for detecting human skin.

We term this camera a "six-band camera". It has two types of 3-CCD cameras: a Vis camera that acquires three images around Vis regions as a conventional color camera and a NIR camera that can acquire three images in individual NIR wavelengths (870, 950, 1050nm). Figure 4 and Figure 5 show a photograph of the six-band camera and a schematic diagram, respectively.

As shown in Figure 5, a NIR reflection mirror is installed to separate the incident light into VIS and NIR light. Each separated light is split into three wavelengths by a prism having three band-pass filter, respectively. Figures 6 and 7 show transmission spectra of the band-pass filters attached in the Vis and NIR cameras, respectively. Developing a camera designed for this configuration, we can obtain different six bands that have the same optical axis.

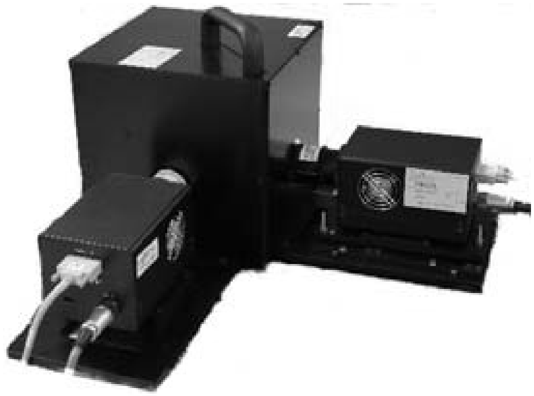

Figure 4. Photograph of six-band camera.

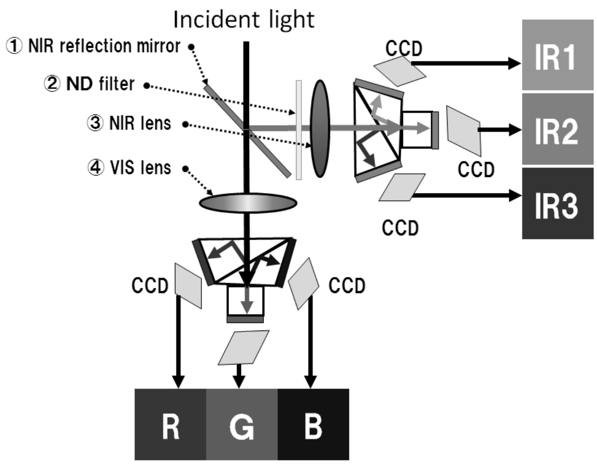

Figure 5. Schematic diagram of the six-band camera.

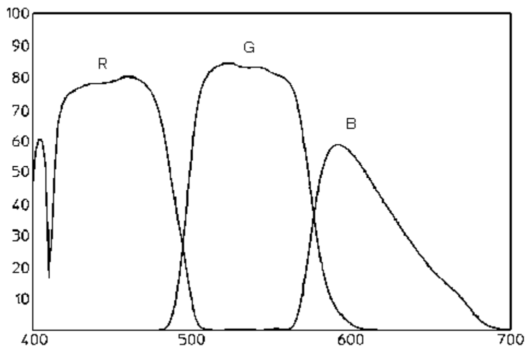

Figure 6. Transmission spectra of band-pass filters for Vis camera.

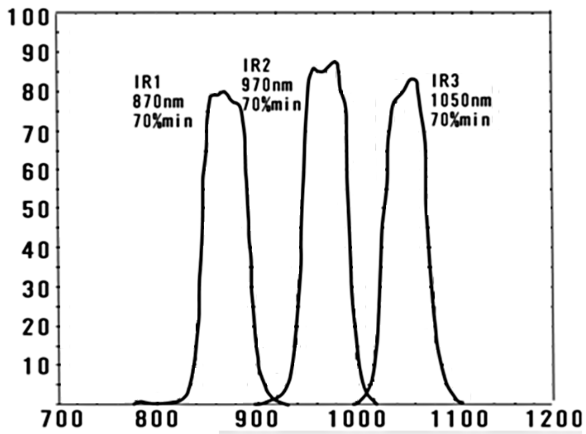

Figure 7. Transmission spectra of band-pass filters for NIR camera.

4 Human Skin Detection

4.1 Algorithm

Conventionally, for detecting human skin from images, visible color image processing methods are used. However, it has been well known that color information is easily affected from a change of illumination condition. On the other hand, [9] proposed a NIR image processing method to detect human skin. This method is based on the characteristics of spectroscopy around NIR region, therefore unaffected from visible light. But, since clouds and water vapor absorbs light with a wavelength of 940nm as shown in Figure 8 and [10], it is difficult to distinguish human skin, which absorbs light with a wavelength of 970nm, from clouds.

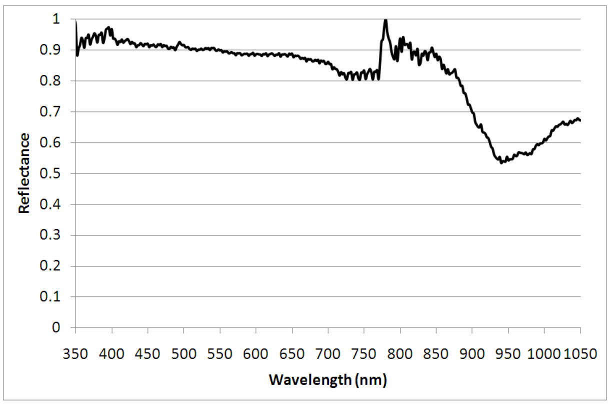

Figure 8. Reflectance spectra of clouds and water vapor.

To overcome above problems, we have developed a six-band camera described in Section II and in addition we propose a method which combines a visible and NIR image processing methods optimized for a six-band camera. In the following section, both visible color and NIR image processing methods for detecting human skin are discussed and then a method to combine both information is proposed.

In the Vis region, we utilize a conventional color image processing method to detect human skin from Vis images. Although many color spaces are proposed, we select the YC_bC_r color space because it has been found to be effective for detecting flesh-colored light [7]. In YC_bC_r, the RGB components are separated into luminance (Y), chrominance blue (C_b), and chrominance red (C_r). Equation (1) is used to convert RGB values into YC_bC_r. We specify a pixel whose color satisfy Equation (2) as a skin [8].

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

where the R, G, and B values are scaled to $[0, 255]$.

$$\begin{cases} 77 \le C_b \le 127 \\ \quad and \\ 133 \le C_r \le 173 \end{cases} \quad (2)$$

In the NIR region, the reflectance of human skin tends to decrease as the wavelength increases from 870nm to 970nm as shown in Figure 2. On the contrary, it increases from 970nm to 1050nm. Thus, we apply a method for detecting human skin based on subtraction in NIR region[9]. Equation (3) is used for detecting human skin in the NIR region.

$$\begin{cases} Lum(870) - Lum(970) > 0 \\ \quad and \\ Lum(1050) - Lum(970) > 0 \end{cases} \quad (3)$$

where $Lum(\lambda)$ expresses the pixel intensity in the λ nm wavelength image.

Finally, both Vis and NIR processing results are combined to distinguish human skin in images. Figure 9 shows a flowchart of the proposed method. Applying this combination, for example, soil whose color is close to skin in the visible color space can be distinguished from human skin because the reflectance of soil is clearly different from one of human skin. In the same way, cloud can be differentiated from human skin by using visible image information.

The proposed method are expected to detect human skin robustly and precisely since it utilizes a wide range of image information (Vis and NIR) acquired by the developed six-band camera and compensates drawbacks in an individual(Vis and NIR) image processing.

4.2 Experimental Results

Figure 10 shows input images: the left side is a Vis image and the right side is a NIR image. The Vis image is expressed by an RGB color model and the NIR

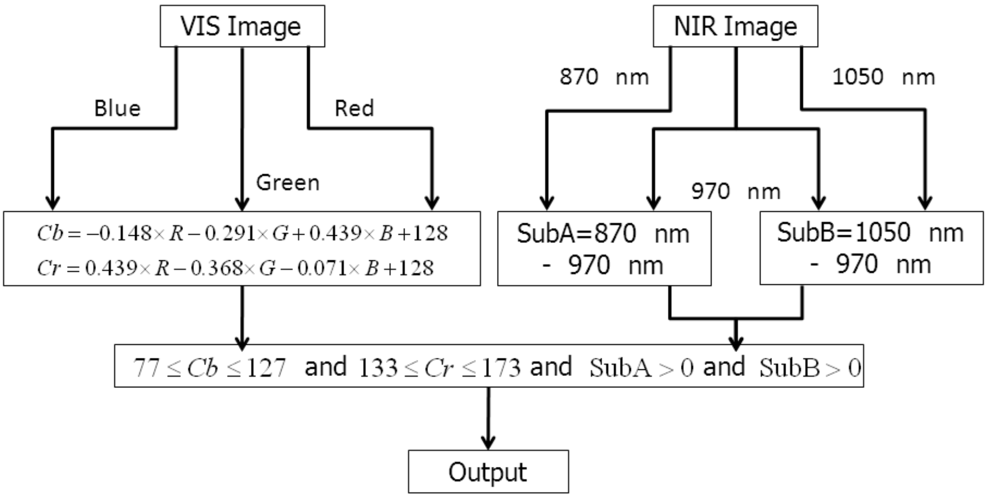

Figure 9. Flowchart of proposed method.

image is expressed by a false color model in which the values of 870, 970, and 1050 nm are displayed as values corresponding to blue, green, and red, respectively. Figure 11 shows images for the results of human skin detection: the left side is the image applying skin detection only for Vis and the right side is a result only for NIR skin detection. Regions detected as human skin appear white in Figure 11.

Figure 10. Input images for Vis/NIR.

Figure 11. Skin detection results (Vis/NIR).

As show in Figure 11, in the Vis region, soil and orange makers are incorrectly identified to human skin. On the other hand, cloud and sky are recognized as human skin. Figure 12 shows the result applying our proposed method for detecting human skin. As shown in Figure 12, by using both VIS and NIR images, we can detect only human skin region successfully.

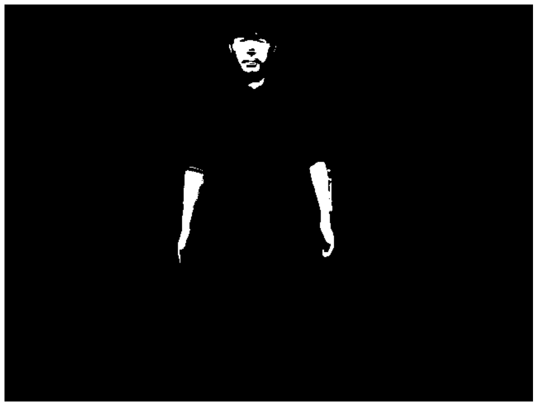

Figure 12. Skin detection result (six bands).

5 Effectiveness of Proposed Method

In this section, we assess the effectiveness of using ADAS to detect pedestrians by comparing the proposed human skin detection method and the conventional pattern recognition method for two cases.

The first case is when a pedestrian is standing on the road. In this experiment, we took an image of a pedestrian standing in the middle of the road. Figure 13 shows the detection results: the left image shows the pattern recognition result and the right image shows the result obtained using the proposed method for detecting human skin.

Figure 13. Pedestrian on a road.

Although, in this case, each method can detect pedestrian correctly, combination of both results is expected to improve performance of pedestrian detection.

The second case is when a pedestrian is occluded by another object. In this experiment, we took images of a pedestrian standing behind a vehicle parked on the side of the road and a pedestrian in the vehicle. The left images of Figure 14 and Figure 15 show the pattern recognition result and the right images show the result obtained using the proposed method.

Figure 14. Pedestrian occluded by a vehicle.

Figure 15. Pedestrian in a vehicle.

As shown in Figure 14 and Figure 15, the pattern recognition method cannot detect the pedestrian in this case. This is because the pedestrian is occluded by ohter objects. On the other hand, the proposed method can succeed to detect a pedestrian based on human skin detection. Compared to pattern based pedestrian detection, our proposed method has advantages as follows; our method can detect occluded pedestrians as long as human skin is observed and requires quite a slight computational cost (see Equation (2) and (3)) although a pattern recognition method generally exhausts it.

6 Conclution

In this paper, we proposed a human skin detection method by considering the reflection characteristics in the Vis and NIR regions. It uses a six-band camera that can acquire both three Vis spectrum images and three other NIR spectrum images. By using such optical absorbance between the different wavelengths, segmentation of materials such as skin can be performed better than using Vis light or NIR light alone. In particular, our proposed method can detect an occluded pedestrian, which a pattern recognition method cannot detect. Therefore, our proposed method is confirmed to be effective.

In order to confirm the feasibility of our theory, we developed a simple human skin detection algorithm based on the analysis from spectroscopy in this paper. For the future work, to recognize human skin more precisely, we will apply an image segmentation method[11][12] to the spectroscopy based image processing. In addition, we will propose a new image features, like Color SIFT[13], employing a six-band color information. Eventually, incorporating with our proposed methods, we aim to develop an advanced pedestrian detection system for realizing vehicles much safer.

References

[1] Karel A. Brookhuis, et al.: "Behavioural impacts of Advanced Driver Assistance Systems: an overview," *European Journal of Transport and Infrastructure Research*, vol. 1, no. 3, pp. 245–253, 2001.

[2] Akihiro Watanabe, et al.: "Lane detection for a Steering Assistance System," *Proceedings of the IEEE Intelligent Vehicles Symposium*, pp. 159–164, 2005.

[3] Yoshikatsu Kimura, et al.: "Stereo vision for obstacle detection," *Proceedings of 13th World Congress on Intelligent Transport Systems*, no. 1082, 2006.

[4] Mineki Soga, et al.: "Pedestrian Detection for a Near Infrared Imaging System," *Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems*, pp. 1167–1172, 2008.

[5] Cao Hui, et al.: "Feature Interaction Descriptor for Pedestrian Detection," *IEICE Transactions on Information and Systems*, no. E93.D(9), pp. 2656–2659, 2010

[6] Dudley Williams: "Methods of Experimental Physics Volume 13," *Academic Press*, 1976.

[7] Rafael C. Gonzalez: "Digital Image Processing," *Prentice Hall*, 2002.

[8] Douglas Chai, et al.: "Face Segmentation Using Skin-Color Map in Videophone Applications," *Proceedings of IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 4, pp. 551–564, 1999.

[9] Shohei Moriakwa, et al.: "Decision Method of the Material Characteristics by Using Three Wavelength Images," *Proceedings of Sixteenth Korea-Japan Joint Workshop on Frontiers of Computer Vision*, pp. 362–367, 2010.

[10] Heidi Cattaneo, et al.: "VCSEL based detection of water vapor near 940 nm," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 60, no. 14, pp. 3269–3275, 2004.

[11] J. Shotton, J. Winn, C. Rother, and A. Criminisi: "TextonBoost: Joint Appearance, Shape and Context Modeling for Multi-Class Object Recognition and Segmentation," in ECCV, pp. 1–15, 2006.

[12] Yousun Kang, Koichiro Yamaguchi, Takashi Naito, and Yoshiki Ninomiya: "Road Scene Labeling Using SfM Module and 3D Bag of Textons," *Proceedings of International IEEE Workshop on 3D Representation for Recognition*, in association with ICCV2009, 2009.

[13] G. J. Burghouts and J. M. Geusebroek: "Performance evaluation of local colour invariants," *Computer Vision and Image Understanding*, vol. 113, pp. 48–62, 2009.