Illumination Invariant Face and Pupil Detection Using a Multiband Camera System

Yousun Kang[1]*, Yoshinori Arai[1], Fumitaka Ono[1], and Duk Shin[2]
[1]Department of Applied Computer Science, Tokyo Polytechnic University
yskang@cs.t-kougei.ac.jp, arai@cs.t-kougei.ac.jp, ono@cs.t-kougei.ac.jp
[2]Department of Electronics and Mechatronics, Tokyo Polytechnic University
d.shin@em.t-kougei.ac.jp

Abstract

Personal video surveillance systems have been developed in the intelligent security systems and monitoring research areas. Computer vision technologies are key to the implementation of surveillance and monitoring systems, applying methods such as face recognition, pupil detection, and human motion analysis. One of the most important techniques is robust illumination invariant object detection for surveillance systems. This paper presents illumination invariant face and pupil detection using a multiband camera for a personal surveillance system. The multiband camera provides images in four bands, near infrared and the three RGB color bands. We introduce a multiband camera system that utilizes two different images, i.e., a color image extracted from a Bayer filter and a near infrared image. We employed a Viola-Jones Face Detector using Haar-like features and AdaBoosting classification for face detection. We also proposed a pupil detection method using a differential of red-band and near-infrared images. Results showed that the proposed multiband camera system is effective for implementing an illumination invariant object detection module.

keywords : Personal surveillance system, Illumination invariant, Face and pupil detection, Multiband camera, Near-infrared image

1 Introduction

Various applications using camera systems have been developed and deployed commercially to improve our daily lives. The performance of camera systems is mainly dependent on image quality and illumination conditions. In the last few decades, we witnessed the development of various camera systems and automation technologies for surveillance. Near-infrared cameras in particular have been utilized in computer vision and Intelligent Transport Systems (ITS) field applications such as remote sensing of vegetation [3], illumination invariant face recognition [7], pedestrian detection [1], and night vision [10].

Special cameras have been used for driving assistance systems such as driver support system [9]. A multiband camera was recently developed for the purpose of integrating a color camera and a near-infrared camera. An input image of the multiband camera, which is called a multiband image, is available in four bands, near-infrared and the three RGB color bands. Multiband cameras are divided into two types, one being a Bayer filter type and the other a prism-based type. Of the two types, Kidono *et al.* [6][5] developed a Bayer filter type directly, which offered an improvement over single-chip digital color cameras. On the other hand, the prism-based type offers clearly separated multi-spectral images in both visible and near-infrared portions of the spectrum. Du *et al.* [2] proposed a prism-based system for capturing multi-spectral videos.

IT CoNvergence PRActice (INPRA), volume: 5, number: 2 (June 2017), pp. 1-10

*Corresponding author: Department of Applied Computer Science, Tokyo Polytechnic University, 1583 Iiyama, Atsugi, Kanagawa, 243-0297, Japan, Tel:+81-46-242-9524

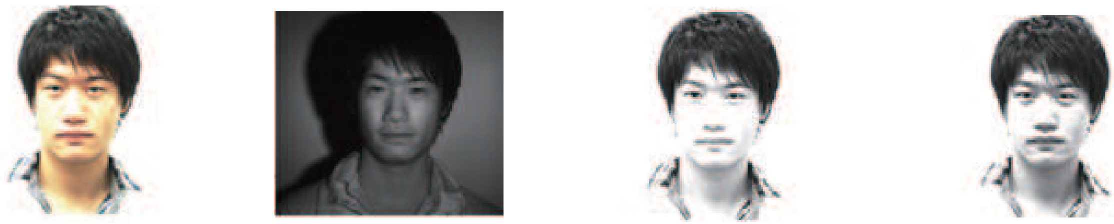


Figure 1: Example face images using a multiband camera: (a) color image (b) near-infrared image (c) visible red band image (d) visible green band image. Visible color is familiar to humans, but near-infrared is efficient for face detection. Exploiting both band types, our multiband camera system provides a more flexible approach.

A multiband camera system can simultaneously obtain three visible wavelength bands and the near-infrared band. The multiband camera can provide various images, i.e. a color image, an infrared image and a visible monochrome image, depending on the requirements of the system. Figure 1 shows face images for comparison between color visible, grey-level visible, and near-infrared. All four images were obtained using the same camera, field of view, and optical axes. The multiband camera system readily allows for processing of images in both visible and infrared bands, and thereby provides greater application flexibility.

The illumination invariant module is important to implement a personal video surveillance or monitoring system. The facial expression monitoring is based on robust illumination invariant face recognition technology, and eye gaze tracking is related to illumination invariant pupil detection modules. Pupil detection in particular is applied in ITS for drowsy driving prevention [8]. In this paper, we present an illumination invariant face and pupil detection module for a personal video surveillance or monitoring system using a multiband camera. We explain the properties of infrared images and structure of the multiband camera in section 2, the illumination invariant face detection in section 3, and the pupil detection module in section 4. We then provide our experimental results showing the effectiveness of the multiband camera system and our conclusions regarding the presented work in section 5.

2 Multiband Camera System

2.1 Infrared Images

The infrared wavelengths on the electromagnetic spectrum roughly lie between 0.7 to 12 microns, in contrast to wavelengths in the visible spectrum, which lie between 0.4 to 0.7 microns. Infrared is divided into near, far, and mid infrared according to their wavelengths as shown in Fig. 2. In the far-infrared domain, the image of an object can depict its temperature emanating heat, so far-infrared cameras are suitable for detection of warmer (or colder) objects. Usually, far-infrared images lack textures and colors.

Another interesting technology is near-infrared. Near-infrared is defined by water absorption, and the effect is formed by strongly reflecting off exterior layers and foliage, such as cloth, skin, tree, leaves, and grass. Objects reflect incident near-infrared radiation very similarly to visible light, and the appearance of objects depends on how the material reflects the near-infrared radiation [1].

In far-infrared images, there are some weak points to implement semantic segmentation and object recognitions. For example, a far-infrared image shows blurred and textureless objects and the far-infrared cameras have poor resolution and are high in cost. In addition, the cameras have difficulty obtaining the

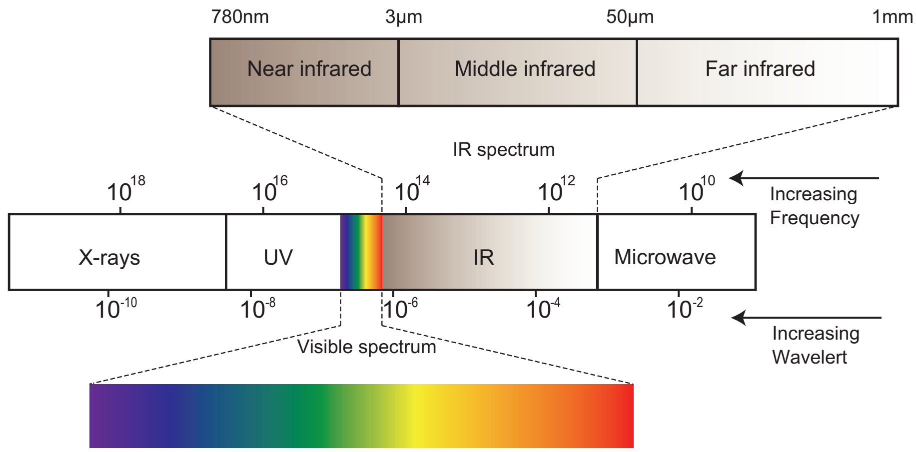


Figure 2: Electromagnetic spectrum with light

same field of view and optical axes as color cameras. However, near-infrared cameras generally provide higher resolutions than far-infrared cameras and are cheaper and smaller. For the purpose of integrating color and near-infrared image wavelengths, multiband cameras are now utilized in computer vision and ITS applications. We introduce the structure of the multiband camera in the next section.

2.2 Multiband camera

This section introduces the structure of a multiband camera system. The prism-based multiband camera provides good reproduction of both visible and near-infrared scenes at a given point in time. The prism-based type is also low cost and is easy to set up, though miniaturization is a challenge for multiband cameras. In this study, we use the prism-based type of multiband camera due to its good reproduction and low cost.

Figure 3 (a) shows the structure of a prism-based multiband camera and how it divides color and near-infrared. The multiband camera is based on a dichroic prism, allowing precise separation of the visible and near-infrared parts of the spectrum. The dichroic prism incorporated in the multiband camera separates the visible part of the spectrum into a wavelength band from 400nm to 650nm (Channel 1) and the near-infrared part into a band ranging from 760 nm to 1000 nm (Channel 2).

The visible channel is referred to as Channel 1 and the near-infrared channel is referred to as Channel 2. Channel 1 and 2 can be configured to operate separately or synchronously. When operating separately, each channel can be triggered and read out independently. In video sequence from the multiband camera, RGB information of the normal color Bayer filter is included in Channel 1. Channel 2 of the video sequence has near-infrared image information. We can separate the two channel images to obtain visible color and near-infrared simultaneously. After a demosaicing [4] of the monochrome image extracted from the Bayer filter in Channel 1, we can obtain the three bands of color from the multiband camera.

3 Illumination Invariant Face Detection

Biometric authentication systems refers to the identification of humans by their characteristics such as fingerprint, voice, face, veins, and iris. They are used to identify individuals in groups that are under surveillance. Among various sources of human characteristics, faces are often used for personal surveillance systems. The pattern of specific organs, such as the eyes, or parts of them, is used in biometric

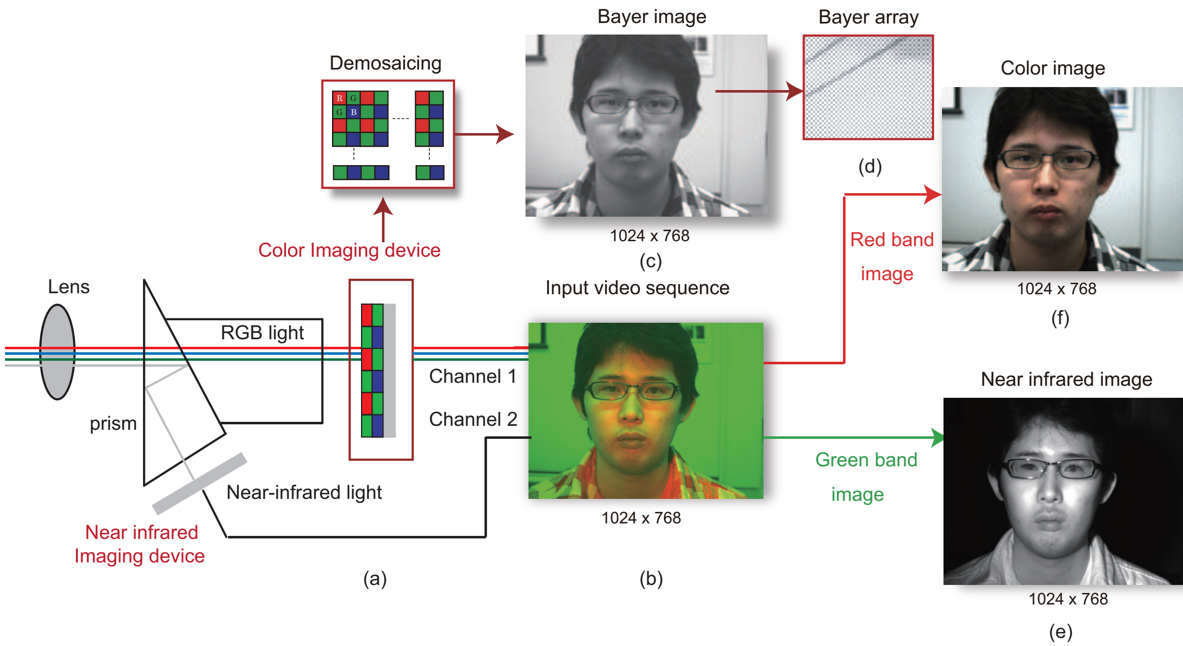


Figure 3: The structure of a prism-based multiband camera. We can obtain both near-infrared and color images from a 2CCD camera. (a) Multiband camera and near-infrared light, (b) input video sequence from multiband camera, (c) Bayer filter image for color, (d) enlarged image pixel for Bayer image and demosaicing process, (e) near-infrared image, (f) color image

identification to uniquely identify individuals. In this section, we present an illumination invariant face detection using a multiband camera system.

The multiband camera is a 2CCD camera (AD-080CL) from JAI Inc. and a near-infrared illuminator includes 140 infrared LEDs with 850 nm. In visible light, a color image is used to detect a face that is familiar to a person's eyes. On the other hand, visible light sometimes creates stress and will be an obstacle in dark places or at nighttime. Thus, depending on the requirements of the user, our system can provide various multiband images. For a face detection algorithm, we employed a Viola-Jones Face Detector [11] using Haar-like features and AdaBoosting classification.

Figure 4 shows the results of a face detection using the multiband camera. In the case of using visible light, we can detect a face using both color and near infrared images as shown in Fig. 4(a). The level of light provides enough illumination, even though at night, as shown in Fig. 4(b). Moreover, our system is able to implement the illumination invariant face detection. Since our multiband camera system can utilize both color and near infrared images, our face detection is not influenced by ambient light as shown in Fig. 4(c). The effectiveness of our face detection system can also be seen in the video sequences as shown in Fig. 5(a) and (b). Since our multiband camera system can utilize both color and near-infrared images, our face detection is not influenced by ambient light.

4 Illumination Invariant Pupil Detection

In this section, we explain the illumination invariant pupil detection module using multiband images. The algorithm procedure of image processing for the illumination invariant pupil detection is as shown in Fig. 6, and we summarized sequentially the module. We acquired video sequences from the multiband camera with a sampling rate of 30 fps. Firstly, we performed face detection from the near-infrared images

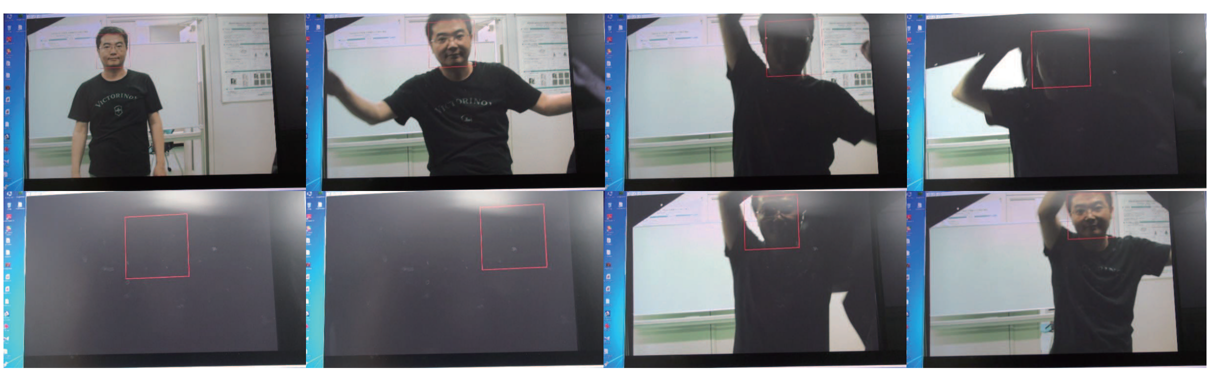
(a)

(b)

(c)

Figure 4: Results of the face detection system: (a) day time (color image at left and near infrared image at right), (b) night time (color image at left and near infrared image at right), (c) illumination invariant face detection (color images)

(a)



(b)

Figure 5: Captured images from video sequences of two results (a) and (b)

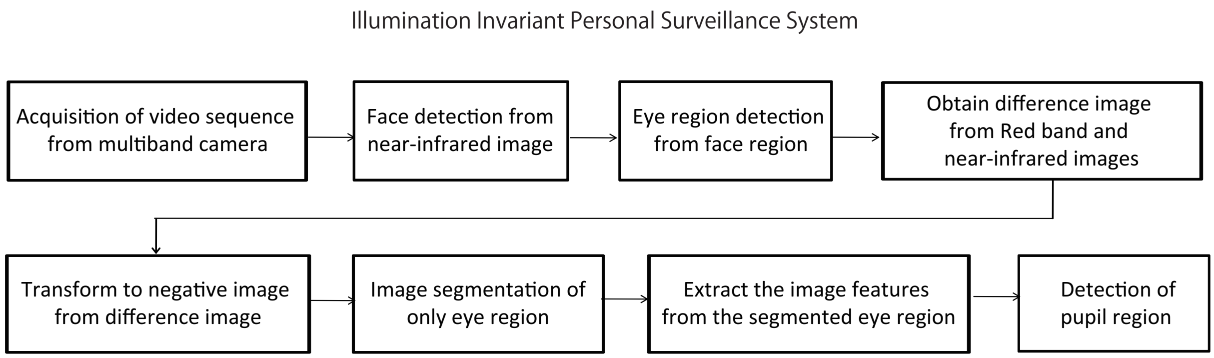


Figure 6: Algorithm of the proposed illumination invariant pupil detection

(image size: 1280*748) and identified eye region from the estimated face region. Secondly, we extracted differential images between red-band and near-infrared images. We then executed image segmentation of eye region with the negative images of the differential images. Thirdly, we extracted the various image features from the segmented eye region. Finally, we registered the final candidate region as a pupil region.

Fig. 7 shows the sample images according to each image processing procedure. The result images on the left side of Fig. 7 were obtained under a bright illumination source. The result images on the right side of Fig. 7 were obtained from a lightless light. Fig. 7 (a) and (i) show the video sequences of multiband images extracted from the multiband camera. The video sequences are separated into color images and near-infrared images. Fig. 7 (b) and (j) are the final result images, which have candidate pupil regions for pupil detection. The yellow boxes in result images (b) and (j) show the detected pupil region.

Fig. 7 (c) and (k) show the near-infrared input image, (d) and (l) show only the red band image from RGB color images. Fig. 7 (e) and (m) show the face region extracted from a near-infrared image. We employed the face detection algorithm by Viola and Jones[11]. Fig. 7 (f) and (n) show the detected eye region. Fig. 7 (g) and (o) show negative images of the differential image.

The differential image is obtained by image subtraction between the red band image and the near-infrared image. The eye region of the negative image is segmented and regarded as the pupil candidate. The region can be represented by various image features such as area, perimeter, centroid, connection component, and circular shape of the region. Finally, we can detect the final pupil region by comparing the prototype of various candidate regions as shown in Fig. 7 (h) and (p).

Fig. 9 shows the result images of the color for illumination invariant pupil detection. In the case of using visible light, we can detect a pupil with gaze variation using both color and near infrared images as shown in Fig. 9(a). The level of light provides enough illumination even though at night, as shown in Fig. 9(b). In darkness, Fig. 10 shows the result images of near-infrared with gaze variation. Since our multiband camera system can utilize both color and near infrared images, our pupil detection is not influenced by ambient light.

Fig. 8 shows the enlarged images of the eye region for detection of pupil region with illumination (top) and without illumination (bottom). Fig. 8(a) is a red band image and (b) is a near-infrared image. Fig. 8(c) shows the differential images between a red band image and a near-infrared image. Finally, we can detect the final candidate region for pupil detection as shown in Fig. 8(d).

To get the accuracy of face and pupil detection, we extracted the 314 video frames from 10 s of acquisition. The extracted images have various illumination conditions and various gaze positions including 3 s of darkness. We detected the face and pupil with 30 fps. The table of Fig. 11 shows the accuracy of face detection and pupil detection using multiband images.Subject C wore glasses, but subjects A and B

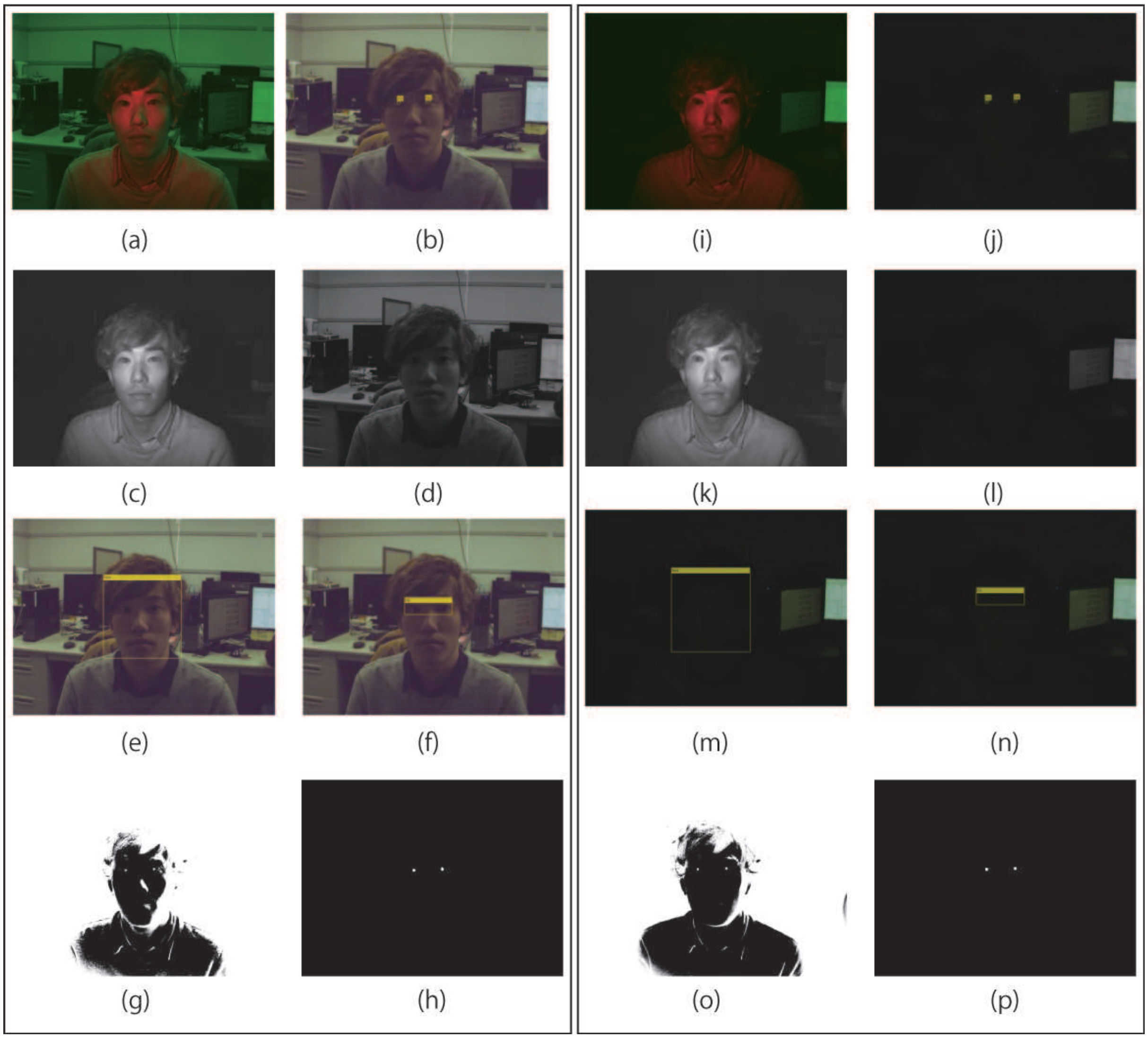
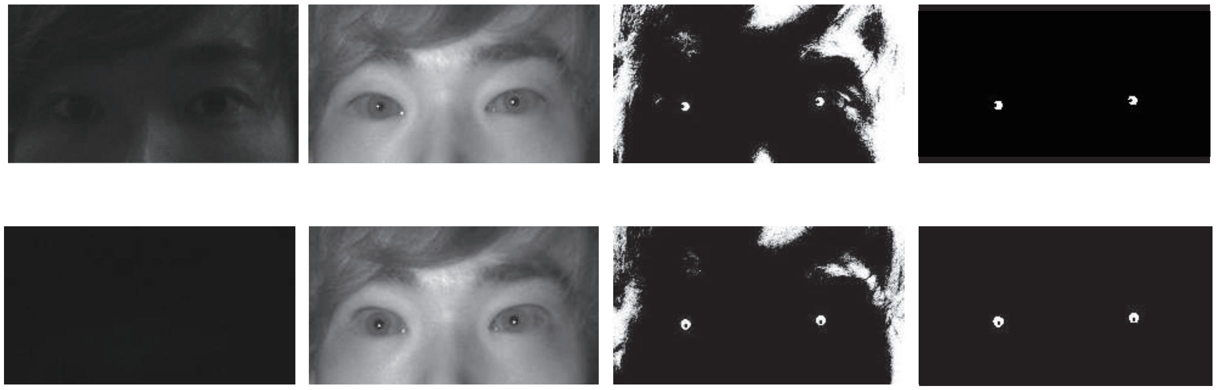


Figure 7: Sample images of pupil detection process with illumination (left) and results images without illumination (right). (a) and (i): A video sequence from the multiband camera, (b) and (j): Result image of pupil detection, (c) and (k): Near-infrared input image, (d) and (l): Red band input image, (e) and (m): Face region, (f) and (n): Eye region, (g) and (o): Negative image of differential image, (h) and (p): Candidate pupil regions

did not. These results show the proposed object detection module gives comparable or better results than the conventional method using only color images when darkness is included. Therefore, we can achieve robust illumination invariant object detection using the multiband camera system.

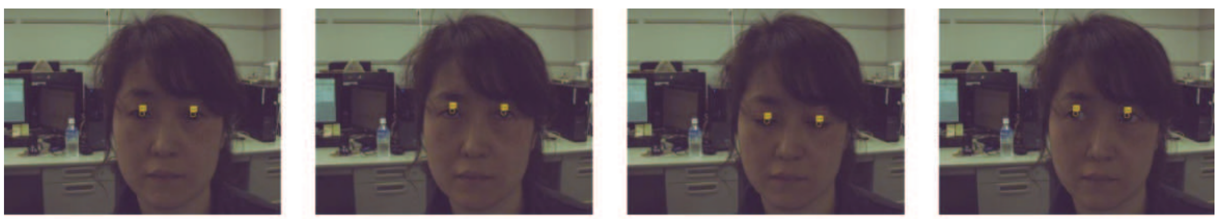
5 Conclusion

This study showed an illumination invariant face and pupil detection using a multiband camera system. We introduced infrared images and explained the structure of the multiband camera system. Through



(a) (b) (c) (d)

Figure 8: Image segmentation result images for pupil detection process with illumination (top) and without illumination (bottom). (a) Red band image, (b) Near-infrared image, (c) Difference image between (a) and (b), (d) candidate regions for pupil detection



(a)

(b)

Figure 9: Sample color images for pupil detection with gaze variation (a) and illumination variation (b)

experimentation, we verified the performance of the illumination invariant application modules with the multiband camera, and confirmed the feasibility of the face detection and pupil detection. The proposed camera system can be applied to various applications such as biometric authentication, personal video surveillance, and autonomous vehicle technology.

Acknowledgments

This work was supported by 'FY2016 MEXT Private University Research Branding Project'.

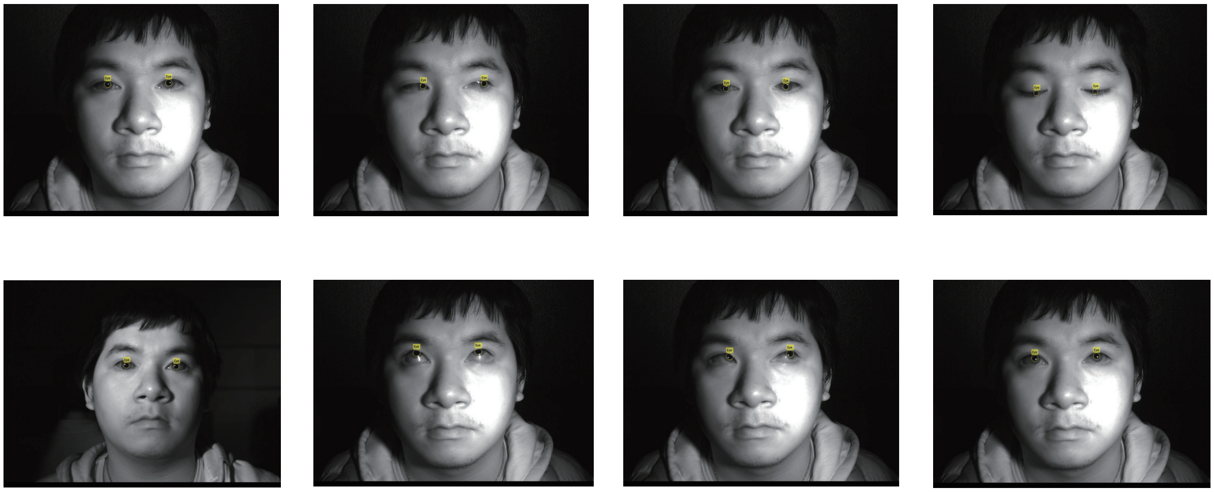


Figure 10: Sample near-infrared images for pupil detection in darkness with gaze variation

		Person A	Person B	Person C	Average
		Multiband image	Multiband image	Multiband image	Multisband image
Face detection rate (Only color image)		80.6 (53.2)	75.4 (49.7)	73.7 (48.6)	76.6 (50.5)
Pupil detection rate	One eye	73.0	71.0	57.5	67.2
	Two eyes	70.4	68.3	49.6	62.8

Figure 11: Accuracy of the face and pupil detection in lightless condition (Subject C wore glasses, but subjects A and B did not.)

References

[1] M. Bertozzi, A. Broggi, A. Fascioli, T. Graf, and M.-M. Meinecke. Pedestrian detection for driver assistance using multiresolution infrared vision. *IEEE Transactions on Vehicular Technology*, 53(6):1666–1678, November 2004.

[2] H. Du, X. Tong, X. Cao, and S. Lin. A prism-based system for multispectral video acquisition. In *2009 IEEE 12th International Conference on Computer Vision*, pages 175–182. IEEE, 2009.

[3] N. Gobron, B. Pinty, and M. Verstraete. Theoretical limits to the estimation of the leaf area index on the basis of visible and near-infrared remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, 35(6):1438–1445, November 1997.

[4] B. K. Gunturk, Y. Altunbasak, and R. Mersereau. Color plane interpolation using alternating projections. *IEEE Transactions on Image Processing*, 11(9):997–1013, September 2002.

[5] Y. Kang, K. Kidono, Y. Kimura, and Y. Ninomiya. Texture-based objects recognition for vehicle environment perception using a multiband camera. In *Proc. of the International Symposium on Visual Computing (ISVC2007), Lake Tahoe, NV, USA, LNCS*, volume 4842, pages 582–591. Springer-Verlag, November 2007.

[6] K. Kidono and Y. Ninomiya. Visibility estimation under night-time conditions using a multiband camera. In *Proc. of the 2007 IEEE Intelligent Vehicles Symposium (IV2007), Istanbul, Turkey*. IEEE, August 2007.

[7] 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(4):627–639, February 2007.

[8] D. Shin, H. S. H, and Y. Uchiyama. Slow eye movement detection can prevent sleep-related accidents effectively in a simulated driving task. *Journal of Sleep Research*, 20(3):416–424, September 2011.

[9] Y. Suzuki, K. Yamamoto, K. Kato, M. Andoh, and S. Kojima. Skin detection by near infrared multi-band for

driver support system. In *Proc. of the 6th Asian Conference on Computer Vision (ACCV), Hyderabad, India, LNCS*, volume 3852, pages 722–731. Springer-Verlag, January 2006.

[10] T. Tsuji, H. Hattori, M. Watanabe, and N. Nagaoka. Development of night-vision system. *IEEE Transactions on Intelligent Transportation Systems*, 3(3):203–209, September 2003.

[11] P. Viola and M. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154, May 2004.

Author Biography

Yousun Kang received the B.S. and M.S. degrees in engineering in 1993 and 1995 from Chosun University in Korea. She received Ph.D. degree in 2010 from Tokyo Institute of Technology, Tokyo, Japan. She is now the professor at Dept. of Applied Computer Science of Tokyo Polytechnic University. Her research interests include robot vision, image processing, texture analysis, and pattern recognition. She is a member of the computer society of IEEE, the Robotics Society of Japan (RSJ), the Institute of Electronics, Information and Communication Engineers (IEICE), and the Institute of Image Electronics Engineers of Japan (IIEEJ).

Yoshinori Arai received the B.S. and M.S. degrees in engineering in 1984 and 1986 from Hosei University in Japan. He received Ph.D. degree in 2010 from Tokyo Institute of Technology, Tokyo, Japan. He is now the associate professor at Dept. of Applied Computer Science of Tokyo Polytechnic University. His research interests include computer vision, image processing, artificial intelligence, and human computer interface. He is a member of the Japanese Society for Artificial Intelligence (JSAI), the Japan Society of Kansei Engineering (JSKE), and IIEEJ.

Fumitaka Ono He has received BE, ME and Ph.D from the University of Tokyo respectively. He is currently Professor Emeritus of Tokyo Polytechnic University, Guest Professor of Tokyo City Univ. and Visiting Researcher of The Univ. of Tokyo. His interested area covers image coding, and image processing including information hiding. He is IEEE Life Fellow, IEICE Fellow and IIEEJ Fellow. He has been engaged in the international standardization work since 1985, and is currently ISO/IEC JTC 1/SC29/WG1 JBIG Rapporteur. He has received Awards from Ministry of Industry and Trade, and also from Ministry of Education and Science, for distinguished achievements on image coding and the contribution for standardization activities.

Duk Shin received the B.S. and M.S. degrees in engineering in 1996 and 1998 from Chosun University in Korea. He received Ph.D. degree in engineering in 2005 from Tokyo Institute of Technology, Tokyo, Japan. He is now the associate professor at Dept. of Electronics and Mechatronics of Tokyo Polytechnic University. His research interests include wearable robot, human centered system, human computer interface, brain machine interface, and bio-signal engineering. He is a member of Society for Neuroscience (SfN), the Japan Neuroscience Society (JNS).