Occupant Detection through Near-Infrared Imaging

Xiaoli Hao[1]*, Houjin Chen[1], Yongyi Yang[2], Chang Yao[1], Heng Yang[1] and Na Yang[1]

[1]*School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, P.R. China*
[2]*Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL 60616, USA*

Abstract

This paper investigates a method to detect and count occupants in vehicles, the purpose of which is to facilitate the task of monitoring and counting of vehicle occupants either by human screeners or by pattern recognition algorithms. The proposed near-infrared (NIR) imaging method can effectively deal with the challenge due to poor light conditions, windshield reflection, tinted windows and shadows on windshields to improve the clarity of the captured image of the vehicle interior. We also proposed an algorithm to process the NIR images. Firstly, the vehicle windshield region was extracted based on optimal edge detections and Hough transform, and ±60° line detector masks and integral projection. Then, the occupants' faces in the region were segmented through AdaBoost-based face detection. Experimental results show that the method has the potential possibility to automatically detect vehicle occupants.

Key Words: Occupant Detection, NIR Imaging, Windshield Reflection, Windshield Extraction, Face Detection

1. Introduction

In recent decades a myriad of technologies including electronic toll collection and license plate recognition have been developed to improve the integrity of enhanced transportation systems. The focus of these technologies has mostly been on the vehicle instead of on the occupants [1]. An occupant detection system can have many practical applications, including: 1) facilitate the operation of high-occupancy vehicle (HOV) lanes, 2) collect data for transportation planning, and 3) monitor vehicles at public facilities, military bases or other sensitive sites. In this paper, we study a method to detect and count the occupants in a vehicle. We use cameras, mounted on lanes and pointed at oncoming vehicles, to capture images of the vehicle interior, from which the vehicle occupants can be counted either by a human screener or by a pattern-recognition algorithm.

*Corresponding author. E-mail: xlhao@bjtu.edu.cn

Some researches [2,3] conducted tests which employed visible range cameras (380 nm–780 nm) to capture the images of vehicle interiors and human screeners to identify occupancy. Low light, windshield reflection (even glare), tinted windows and shadows on windshields compound the problem of viewing occupants in the vehicle compartment [2–4]. To change the poor light conditions (night or overcast skies), we can use illuminator. The challenge of using visible range illuminator to lighten vehicle compartments during nighttime is that the strong visible light could distract the drivers to cause traffic accidents. The Near-Infrared (NIR) light is invisible to human eyes, so there is no risk of distraction to drivers during daytime and nighttime. Ioannis Pavlidis et al. [5] have proposed an occupant detection method which uses two frontal-view of vehicle images taken by irradiating lower band NIR light (1100 nm–1400 nm) and upper band NIR light (1400 nm–1700 nm), respectively. The method utilizes the unique reflectance characteristics of human skin in the NIR spectrum to simplify

the occupants' faces detection. A weak point of the method is that it detects occupants' faces in an image generated by subtraction of the two band NIR images. For example, it is difficult to match the scenes in the two band NIR images, coming from two sensors, at pixel level in a practical imaging system. Besides, the imaging devices of the technique are very expensive, which limits its application. J. W. Wood et al. [6] captured side-views of vehicles by using a NIR camera illuminated by a NIR light source (850 nm–1100 nm), but didn't propose a pattern recognition algorithm to process the side-view vehicle images.

The presence of window glass degrades the image quality of the vehicle interiors (e.g., windshield reflection, tinted windows and shadows on windshields). For example, Figure 1 shows an image, taken by a camera operating in the visible range. Here the vehicle interior becomes barely visible due to the strong reflected light off the windshield. It is impossible to count or identify occupants either by human screeners or computers.

In the former part of this paper, mainly from the angle of eliminating the effect of reflection off the windshield, we investigate a NIR imaging method to capture improved frontal-view images of vehicle interiors. In fact, it can also effectively deal with the challenge due to poor light conditions, tinted windows and shadows on windshield to obtain better images of vehicle interiors. The performance of the method will be given in section 2. The proposed imaging method utilizes a narrow NIR range that is 850 ± 20 nm. In this range, we can use classical CCD sensors and accessories, therefore the imaging devices are reliable, low cost and high-speed. Besides, the images acquired in the NIR range are very similar to visible images, so we can capitalize on state-of-art face detection methods (mainly developed on visible images) to efficiently perform occupants' faces detection.



Figure 1. Occupant visibility is hampered by windshield reflection.

In the latter part of this paper, we present an approach to process the NIR images. Firstly, an optimal edge detection and Hough transform are employed to detect the border lines of the windshield to segment the windshield region. Then, the AdaBoost-based algorithm is used to scan the extracted vehicle windshield region and locate occupants' faces.

The rest of the paper is organized as follows: In section 2 the mechanism of vehicle glass reflection is analyzed and our imaging solution is described and validated with experimental results. In section 3 the image processing algorithm, including extracting vehicle windshield and face detection, is presented. In section 4 the paper is concluded and future work is mentioned.

2. Near-Infrared Imaging Method

2.1 Problem Analysis

Consider the imaging setup in Figure 2. The camera points at the vehicle. When the incident environmental light, which comes from the sun, sky, clouds, the surrounding objects, or the illuminator, reaches the vehicle windshield, one portion of the light is reflected back by the vehicle windshield directly to the camera; let's denote the corresponding intensity of this reflected portion at a camera pixel by I_{glass}. In the mean time, another portion of the incident light penetrates through the windshield and goes into the vehicle.

For simplicity, the refraction of the light at the vehicle windshield is ignored and not shown in Figure 1, which should not affect the analysis.

The occupant's face, which is a diffusing surface, that reflects the surrounding incident light rays in many

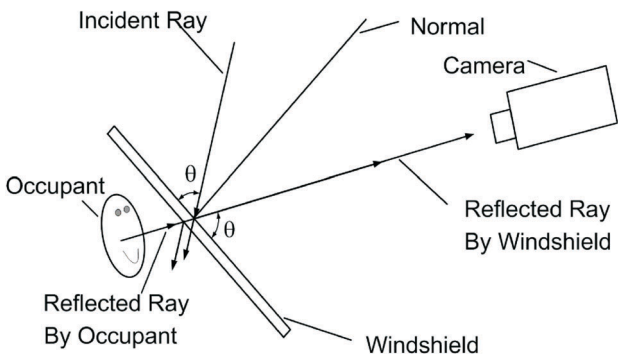


Figure 2. Formation of vehicle windshield reflection in imaging.

different directions. We assume that one reflected ray penetrates through the windshield, and then reaches the same camera pixel along the same optical path. Let's denote the intensity of this portion at the same camera pixel by I_{occupant}.

Therefore, the total light intensity at the camera pixel, denoted by I_{camera}, is given by

$$I_{\text{camera}} = I_{\text{glass}} + I_{\text{occupant}}$$

When I_{glass} is much smaller than I_{occupant}, the vehicle glass reflection is negligible, and the camera can provide a clear image of the vehicle interior. On the other hand, when I_{glass} becomes comparable to or even larger than I_{occupant}, the glass reflection become dominant over the light from the interior of the vehicle, resulting in an image as shown in Figure 1.

In order to reduce the glass reflection entering into the camera ($I_{\text{glass}} \to 0$), one may optimize the positions and view-angles of the camera. However, this is hardly practical, because this optimal setting likely will change with the movement of the sun, and the glass reflection can also result from other surrounding light sources.

2.2 Near-Infrared Imaging Approach

To address the challenges of described above, we adopt the imaging setup as illustrated in Figure 3. The different components of this configuration are explained below.

A camera, mounted above the traffic lane, points in the direction of an oncoming vehicle. This camera operates in the spectral range of 400 nm–1100 nm, which is widely available in lost-cost, commercial products.

The camera is equipped with a narrowband NIR filter (Figure 4) in order to eliminate interference from visible, ultraviolet, and other out-of-band NIR light. Figure 5 shows the spectrum of solar irradiation, which consists of near-infrared, visible, and ultraviolet light.

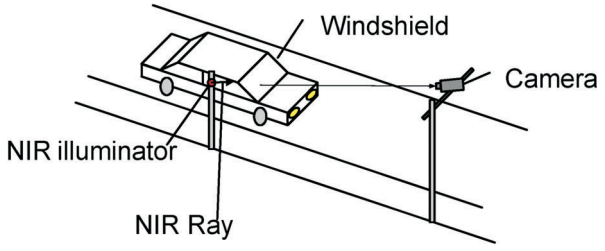


Figure 3. Configuration of NIR imaging.

Matching NIR illuminators (Figure 6), installed at the sides of the vehicle lane, are used to emit strong NIR light into the vehicle window to illuminate the scene of the vehicle interior. With this configuration, the camera is blind to the most of light both emitted by the NIR illuminator and reflected by the side and frontal windows.

When in operation, the NIR illuminator shines light into an approaching vehicle and the camera captures the scene of the vehicle interior through the narrowband filter.

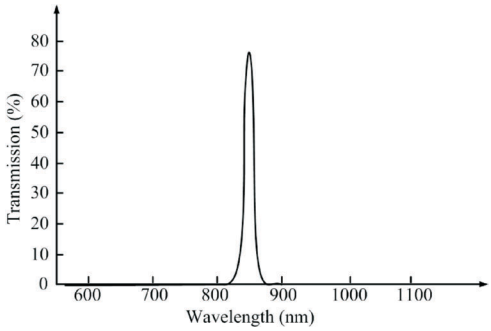


Figure 4. Transmittance spectrum of the NIR filter.

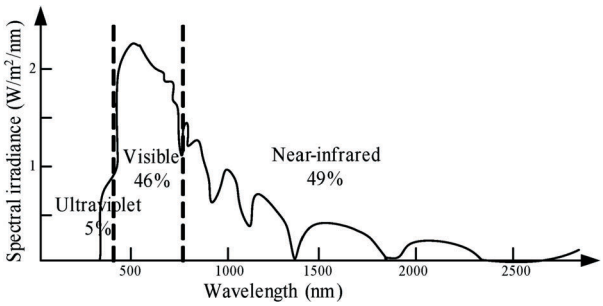


Figure 5. Solar irradiance spectrum at the earth surface.

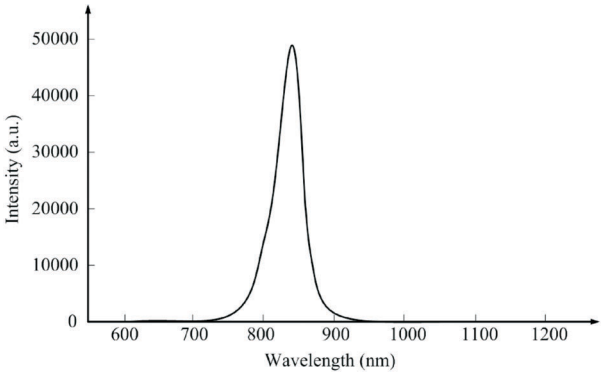


Figure 6. Irradiation spectrum of the NIR illuminator.

The choice of NIR spectrum is out of several considerations. First, the NIR spectrum is effective in penetrating through vehicle windows, including most types of tinted windows (see Figures 7, 8, and 9). Second, there is no risk of distraction to the driver because NIR light is invisible to the human eye. The NIR imaging system can function during both day and night, and even in certain adverse weather conditions.

Below we provide an analysis of the imaging configuration in Figure 3. For convenience, let's consider the diagram in Figure 10, which illustrates the major light paths entering the camera. Since the camera is equipped with the narrowband filter, which passes the infrared light in the narrowband and blocks the other components of light, only the in-band NIR light is shown in Figure 5 for the simplicity.

The total intensity of the light arriving at the point of the camera lens is given by:

$$I_{\text{camera}} = I_{\text{glass}} + I_{\text{occupant_environ}} + I_{\text{occupant_illum}}$$

where I_{glass} is the windshield reflection from the environment, $I_{\text{occupant_environ}}$ is the occupant reflection from the environment, and $I_{\text{occupant_illum}}$ is the occupant reflection from the NIR illuminator.

With the imaging setup in Figure 3, the component I_{glass} is significantly reduced by NIR narrow-band filter. And, by controlling the power of the NIR illuminator, we can achieve $I_{\text{occupant_illum}}$ much larger than I_{glass}, thereby further reducing the effect of reflection captured by the camera. Therefore, according to the equation (2), the total intensity of the light arriving at the camera lens (I_{camera}) is mainly determined by the reflection from occupants ($I_{\text{occupant_environ}}$ and $I_{\text{occupant_illum}}$), not by the reflection from windshield (I_{glass}).

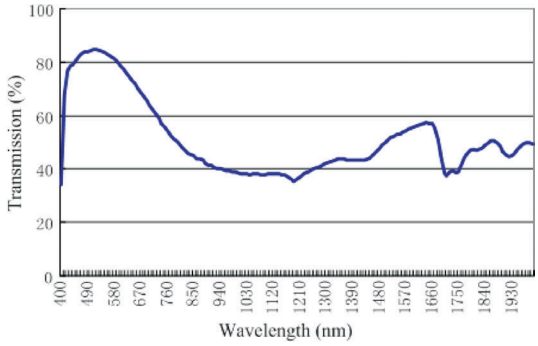


Figure 7. Transmittance of a typical non-tinted windshield.

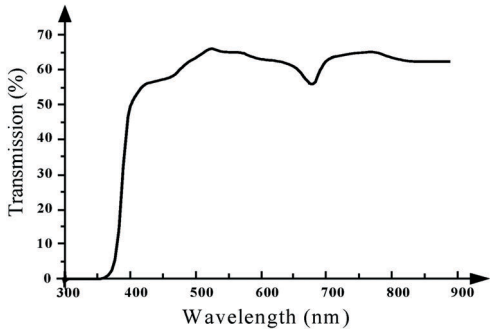


Figure 8. Transmittance of a sample windshield film.

2.3 Experimental Validation

We conducted experiments to validate the proposed NIR imaging method. Figure 11(a) is taken by a color camera which operating in visible band, and Figures 11(b), (c) and (d) are taken by Black/White (B/W) cameras of the same type, which has the spectral range of 400 nm–1100 nm. Figure 11(b) is taken by a B/W camera

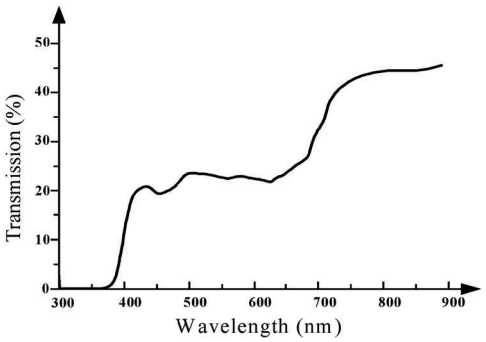


Figure 9. Transmittance of a typical vehicle side-window film.

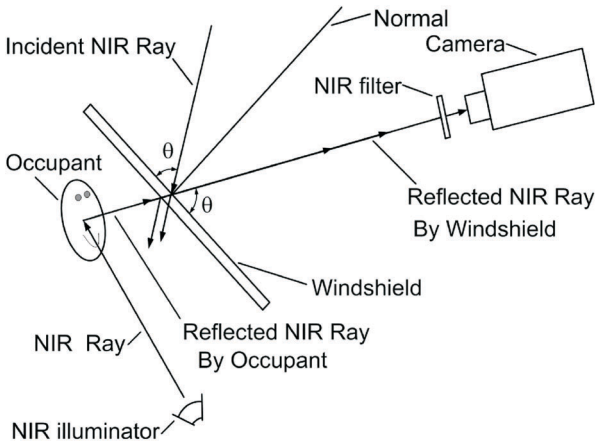


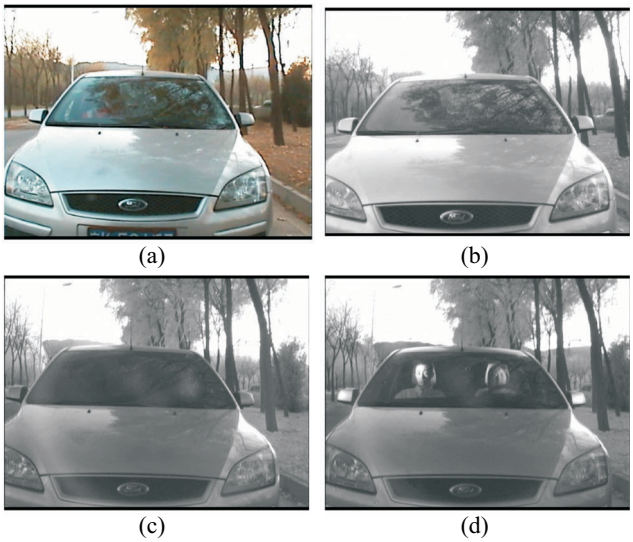
Figure 10. Eliminating windshield reflection with NIR imaging.

operating in 400 nm–1100 nm. Figure 11(c) is taken by a B/W camera with a narrow-band filter of 830 nm–870 nm. Figure 11(d) is taken by adding two matching NIR illuminators, which set on both side of the lane. The vehicles' windows are tinted.

In Figures 11(a), (b) and (c), here the vehicle interior becomes barely visible due to low illumination inside, strong reflected light off the windshield and the trees' shadow. This poses great challenge for the task of counting or identifying occupants either by human screeners or computers. Figure 11(d) is captured based on our NIR imaging method with the imaging configuration in Figures 3, 4 and 6. As can be seen, the windshield reflection and the trees' shadow are largely eliminated, with the occupants clearly visible.

Figure 12 give two sampled vehicle images. One is taken during daytime and the other is taken during nighttime.

From these results, it can be seen that the proposed NIR imaging method could achieve clear images of the vehicle interior, which allow counting of the occupants by human screeners.



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

Figure 11. Comparison of various imaging setup. (a) serious windshield reflection and tree's shadow (taken by a color camera in visible-band). (b) serious windshield reflection and tree's shadow (taken by a black-white camera operating in 400 nm–1100 nm). (c) windshield reflection and tree's shadow (taken by a black-white camera of the same type operating in 830 nm–870 nm). (d) windshield reflection and tree's shadow are eliminated (taken by a black-white camera of the same type operating in 830 nm–870 nm and two matching NIR illuminators on both sides).

3. Occupant Detection

We designed an image processing algorithm to process the NIR images. This is performed by following steps: (1) the extraction of the region of interest (the windshield region); (2) the detection of occupants' faces in the windshield region.

3.1 Vehicle Windshield Extraction

This part faces following challenges: (1) windshields of different vehicle types vary in shape, size and relative position; (2) the contrast between the vehicle windshield region and the vehicle body is greatly different because vehicle windshield regions are usually dark but vehicle color and lightness vary greatly. Although vehicles have different shapes, relative position of windshield, and colors, their front windshields have similar geometric features. The upper and lower borders of the windshield are close to horizontal lines, and the left and right borders are close to ±60° diagonal lines (Car) or vertical lines (trucks).

The extraction of the vehicle windshield involves performing an edge detection and Hough transform to detect the upper and down horizontal lines of the windshield, and employing ±60° line detector masks and integral projection to detect the left and right borders of the windshield.

The part employs optimal edge detection [7] to obtain the edge response of upper and down horizontal lines of the vehicle windshield, which has better performance than common edge operators (Prewitt, Sobel, Kirsch [8], and standard Canny). Canny defined three criteria (detection, localization and single response to a single edge) and used the criteria in numerical optimiza-



(a) (b)

Figure 12. Occupants' faces detection. (a) During daytime. (b) During nighttime.

tion to derive detectors for a class of edges. Under the step edge model, Canny pointed out that there is a natural uncertainty principle between detection and localization performance. Let's denote a step edge $g(x) = Au(x)$, where $u(x)$ is the unit step function and A is the amplitude of the edge. The edge is corrupted by a noise, with average noise amplitude n_0, and the edge operator function $f(x)$ has a finite impulse response bounded by $[-w, w]$. The signal-to-noise ratio (which corresponds to the detection performance) is defined as

$$SNR = \frac{d\left|\int_{-W}^{0} f(x)dx\right|}{n_0\sqrt{\int_{-W}^{W} f^2(x)dx}}$$

The reciprocal of average localization error is

$$Localization = \frac{d|f'(0)|}{n_0\sqrt{\int_{-W}^{W} f'^2(x)dx}}$$

It is desirable to maximize Equations (3) and (4) simultaneously. However, there is a direct trade-off between the two which can be varied by changing the spatial width of the operator. Suppose that we form a spatially scaled operator $f_w(x)$ from $f(x) : f_w(x) = f(x/w)$, the SNR and the localization are

$$\begin{cases} SNR_w = \sqrt{w}\, SNR \\ Localization_w = \dfrac{1}{\sqrt{w}} Localization \end{cases}$$

It can be seen that the larger operators give better detection performance and poorer localization performance.

Canny indicated that a spatial operator elongated along the edge direction can improve both detection and localization to overcome the trade-off limitation. The examples of some elongated operators are shown in Figure13.

Assuming the spatial operator is $f(x, y)$. If we scale it along the direction: $f_w(x, y) = f(x/w, y)$, we obtain

$$\begin{cases} SNR_w = \sqrt{w}\, SNR \\ Localization_w = \sqrt{w} Localization \end{cases}$$

Both performances are improved.

For common edge detection, the width of the edge operator should be limited, so it can detect edge pixels correctly along possibly complicated edge contours [9]. In our application, however, the vehicle windshield has the border of two horizontal lines and two diagonal lines, so it is obviously preferable to use elongated horizontal and diagonal edge masks (see Figure 13) to improve both detection and localization performance.

After extracting the horizontal windshield region, we detect side borders of the windshield by employing ±60° line detector masks and integral projection (see our previous work [10]). At last, the windshield region can be segmented from the vehicle images.

3.2 Occupants' Faces Detection

Most Face detection algorithms were developed on visible images. In our work, the NIR images in the range of 850 ± 20 nm are close enough to the visible light spectrum to capture the structure of the face. Therefore, we can capitalize on some of the state-of-art face detection methods to perform occupants' faces detection on the NIR images.

Numerous presentations have been proposed for face detection, including pixel-based [11], parts-based [12], local edge features [13], Haar wavelets [14] and Haar-like features [15]. Haar-like features. encode the existence of oriented contrasts exhibited by a human face and their special relationships [16]. Using Haar-like features, the AdaBoost-based face detection algorithm by Viola and Jones [15] demonstrated that faces can be fairly reliably detected in real-time under partial occlusion. Extended features can be used with AdaBoost for improving the performance of face detection [17–19]. The extended features in [17], which are employed by this paper, are shown in Figure 14.

Viola and Jones' work [15] includes three key contributions. The first is the use of Haar-like features that are quickly computed using an approach called "integral

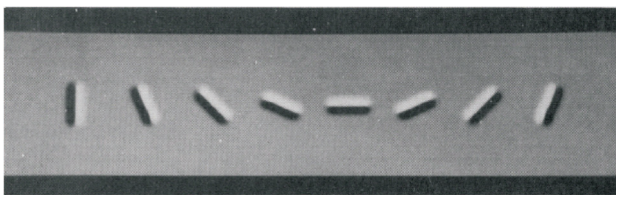


Figure 13. Elongated directional operators.

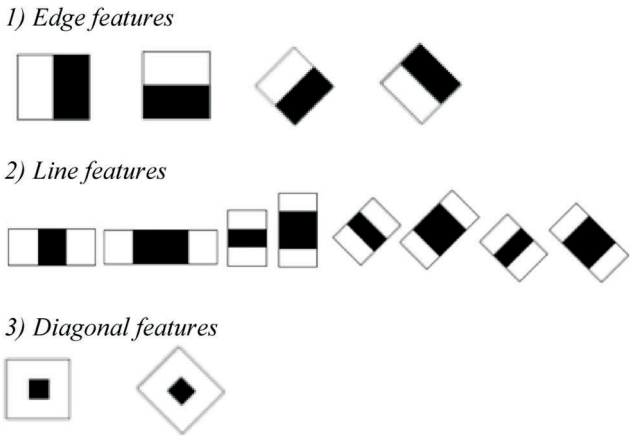


Figure 14. Haar-like features.

image". The second is a learning algorithm, based on AdaBoost, which selects a small number of critical features from the pool of thousand of potential Haar-like features and combines them to synthesize the final strong classifier. Thirdly, the final detector employs a cascade architecture (Figure 15), which combines increasingly more complex classifiers and is aimed to quickly discarded most background regions of the image at the earliest stage possible while spending more computation on face-like regions at later stages. Based on the above AdaBoost-based method, this paper scans through the segmented region of vehicle windshield to locate occupants' faces.

3.3 Experimental Results

The whole algorithm has to extract the vehicle windshield firstly and then detect occupants' faces in the region. Figure 16 visually illustrate the performance comparison between the elongated horizontal operator and the common edge operator. Figure 16(a) is an original NIR vehicle image, where the left half part of the vehicle is lighter than the right part. Figures 16(b)~(f) are the results of several different horizontal operators collecting edge response from the vehicle image. Using the common horizontal operator, Figure 16(b, c, d) not only lost partial lower horizontal border line of the vehicle windshield, but also kept more extraneous information.

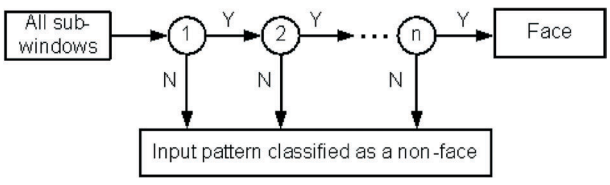


Figure 15. Cascade of classifiers with n stages.

For convenient comparison, we modified the standard Canny edge detection algorithm to only response horizontal edges, as shown in Figure 16(e). Most of extraneous pixels are eliminated, but the lower horizontal border line of the windshield remains half. We designed 5*11 elongated horizontal operators and correctly obtained the upper and lower border lines of the windshield, as shown in Figure 16(f). More experimental results also show that the optimal directional operators are more robust under varied illumination conditions.

By using the optimal edge detection and Hough transform, we conduct the windshield extraction experiment on 182 NIR vehicle images. The correct detection rate (CDR) is 92.8%.

After the windshield region was extracted from the original NIR images, the AdaBoost-based algorithm scans the region and locates occupants' faces. An example result is shown in Figure 17. The performance of the face detection greatly depends on the quality of NIR images. During clear day, there is strong NIR component in sunlight (see Figure 5), and lighting conditions vary greatly as weather and time change, which degrade the quality of images. During cloudy weather and nighttime there is no strong NIR light in surroundings, we can obtain more clear and stable image signals. Besides, tinted

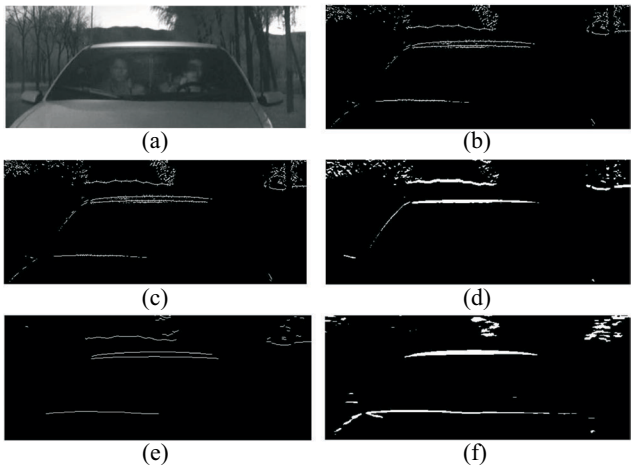


Figure 16. Elongated directional operator vs. common edge detection operators. (a) Original image. (b) Prewitt horizontal operator. (c) Sobel horizontal operator. (d) Kirsch 3*3 horizontal operator. (e) Modified Canny horizontal operator. (f) Elongated horizontal operator.



(a) (b)

Figure 17. An example result of Occupants' faces detection. (a) Original image. (b) segmented windshield region and located faces.

windows of different types and non-tinted windows have significant differences in transmittance of NIR light, which also cause non-stable quality of vehicle interior images. In order to maintain a stable quality of images, it is necessary to design a sophisticated NIR illuminator which will adaptively adjust illumination to keep an optimal lighting level in the scene of the vehicle windshield region. This is the further research direction of this work.

4. Conclusion

In this paper, we investigated a NIR imaging method which can effectively eliminate vehicle glass reflection and obtain improved images of vehicle occupants. The proposed NIR method can function during both day and night, and has the advantages of being reliable, low cost, and high-speed. Moreover, we developed an algorithm to detect the occupants in the NIR images. Firstly, we employed optimal elongated directional operators and Hough transform to segment the windshield region. Then, we conduct the AdaBoost-based face detector to locate the occupants' faces in the windshield region. Experimental results show that the method has the potential possibility to detect vehicle occupants. More future work is needed for improving the adaptive performance of NIR illuminator to keep a stable quality of image signals under varied conditions of weather and tinted/non-tinted windows, and for investigating more robust occupant detection algorithms.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (No. 60972093, No. 60872081); Beijing Natural Science Foundation (No. 4092030).

References

[1] John Wikander, *Automated Vehicle Occupancy Technologies Study*, USA: Texas Transportation Institute, Aug. (2007).

[2] John Billheimer, Ken Kaylor and Charles Shade, *Use of Videotape in HOV Lane Surveillance and Enforcement: Final Report*, U.S. Department of Transportation, Sacramento, California, March (1990).

[3] Shawn Turner, *Video Enforcement of High Occupancy Vehicle Lanes: Field Test Results for I30 in Dallas*, Transportation Research Record No. 1682, Transportation Research Board, Washington, D.C., pp. 26–37 (1999).

[4] Albert Gan, Rax Jung, Kaiyu Liu, Xin Li and Diego Sandoval, *Vehicle Occupancy Data Collection Methods*, Florida International University for Florida Department of Transportation, Feb. (2005).

[5] Ioannis Pavlidis, Vassilios Morellas and Nikolaos Papanikolopoulos, "A Vehicle Occupant Counting System based on Near-Infrared Phenomenology and Fuzzy Neural Classification," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 1, pp. 72–85 (2000).

[6] Wood, J. W., Gimmestad, G. G. and Roberts, D. W., "Covert Camera for Screening of Vehicle Interiors and HOV Enforcement," *Proc. SPIE - The International Society for Optical Engineering*, Vol. 5071, pp. 411–420 (2003).

[7] John Canny, "A Computational Approach to Edge Detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 8, pp. 679–698 (1986).

[8] Kirsch, R., "Computer Determination of the Constituent Structure of Biological Images," *Computer Biomedical Research*, Vol. 4, pp. 315–328 (1971).

[9] Moo, H., Chellappa, R. and Rosenfeld, A., "Performance Analysis of a Simple Vehicle Detection Algorithm," *Image and Vision Computing*, Vol. 20, pp. 1–13 (2002).

[10] Hao, X., Chen, H., Wang, C. and Yao, C., "Occupant Detection through Near-Infrared Imaging," *Cross-Strait Conference on Infromation Science and Technology*, Qinhuangdao,China, June, pp. 332–335 (2010).

[11] Rowley, H., Baluja, S. and Kanade, T., "Neural Network-Based Face Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20,

pp. 23–38 (1998).

[12] Heisele, B., Serre, T. and Poggio, T., "A Component-Based Framework for Face Detection and Identification," *International Journal of Computer Vision*, Vol. 74, pp. 167–181 (2007).

[13] Fleuret, F. and Geman, D., "Coarse-to-Fine Face Detection," *International Journal of Computer Vision*, Vol. 41, pp. 85–107 (2001).

[14] Schneiderman, H. and Kanade, T., "Object Detection Using the Statistics of Parts," *International Journal of Computer Vision*, Vol. 56, pp. 151–177 (2004).

[15] Viola, P. and Jones, M., "Rapid Object Detecting Using a Boosted Cascade of Simple Features," *Proc. CVPR*, pp. 511–518 (2001).

[16] Yang, M.-H., "Face Detection," *Encyclopedia of Biometrics*, Part 6, pp. 303–308 (2009).

[17] Lienhart, R., "An Extended Set of Haar-Like Features for Rapid Object Detection," *IEEE ICIP 2002*, Vol. 1, pp. 900–903 (2002).

[18] Messom, C. H. and Barczak, A. L. C., "Fast and Efficient Rotated Haar-Like Features Using Rotated Integral Images," *Australian Conference on Robotics and Automation ACRA*, pp. 1–6 (2006).

[19] Abiantun, R. and Savvides, M., "Boosted Multi-Image Features for Improved Face Detection," *IEEE Applied Imagery Pattern Recognition (AIPR) Workshop*, pp. 1–8 (2008).

Manuscript Received: Dec. 17, 2010
Accepted: Feb. 15, 2011