A STUDY OF NEAR-INFRARED (NIR) FILTER FOR SURVEILLANCE APPLICATION

Mohd Farid Mohd Ariff*, Mohammad Ehsan Kosnan, Zulkepli Majid, Albert K Chong, Khairulnizam M Idris

Photogrammetry & Laser Scanning Research Group, Faculty of Geoinformation and Real Estate, Universiti Teknologi Malaysia, 81310, UTM Johor Bahru, Johor, Malaysia.

Graphical abstract

Abstract

Lately, most illegal activities occur in the dead of night when most of the surveillance cameras cannot capture movements clearly. Therefore, Near-Infrared (NIR) filter was used to increase the visualization of suspect identification when the image or footage is captured in a dark environment. The objective of this study was to determine the optimum NIR filters based on the stability of the camera calibration parameters and to evaluate the accuracy of mapping. In this study, four NIR filters with different wavelengths (715, 780, 830, and 850 nm) were tested. The investigation comprised: (1) the calibrations of the camera and NIR filters and (2) a case study involving a simulation test for surveillance application. The type of sensor used was a digital video camera (Sony HC5E HDV) and the camera was set up at multiple stations to form a single convergence configuration. The statistical Analysis of Variance (ANOVA) was used in this study to find (differences in) the significance of the NIR imaging in the calibration and three-dimensional (3D) measurement. The results showed that the camera parameters varied for every type of filter used and this influenced the 3D measurement of the object mapping. In summary, the 850 nm NIR filter was the most optimum for surveillance application based on the stability of the camera calibration and the standard deviation in the mapping accuracy.

Keywords: Near-infrared filters, camera calibration, three-dimensional (3D) measurement

Abstrak

Sejak kebelakangan ini, kebanyakan aktiviti-aktiviti jenayah bertak pada waktu lewat malam di mana kebanyakan kamera pengawasan gagal merakam dengan jelas. Oleh yang demikian, teknologi infra-merah dekat telah digunakan bagi meningkatkan visualisasi identiti suspek apabila imej atau video dirakam dalam persekitaran yang gelap. Objektif kajian ini adalah untuk menentukan panjang gelombang Inframerah dekat yang optimum berdasarkan kestabilan parameter kalibrasi kamera dan menilai ketepatan pemetaan. Dalam kajian ini, empat penapis yang berbeza panjang gelombang (715, 780, 830 & 850 nm) telah diuji. Kajian ini meliputi: (1) kalibrasi kamera dan penapis infra-merah dekat dan (2) kajian kes yang melibatkan ujian simulasi untuk aplikasi pengawasan. Jenis penderia yang digunakan ialah kamera video berdigit (Sony HC5E HDV) dan kamera dipasang-siap di berbilang stesen bagi membentuk konfigurasi menumpu. Analisis statistik varians (ANOVA) telah digunakan dalam kajian ini untuk mencari perbezaan signifikan pengimajaan infra-merah dekat dalam kalibrasi dan pengukuran tiga dimensi (3D). Keputusan menunjukkan bahawa parameter kamera berbeza bagi setiap jenis penapis yang digunakan dan ini mempengaruhi pengukuran 3D bagi pemetaan objek. Secara ringkas, penapis infra-merah dekat 850 nm adalah yang paling optimum untuk aplikasi pengawasan berdasarkan kepada kestabilan kalibrasi kamera dan sishan piawai dalam ketepatan pemetaan.

Kata kunci: Penapis infra-merah dekat, kalibrasi kamera, pengukuran tiga dimensi (3D)

© 2015 Penerbit UTM Press. All rights reserved
1.0 INTRODUCTION

Close-Range Photogrammetry (CRP) has been widely used in fields other than for surveying purpose. Nowadays, this technology is frequently implemented because the resolution or ground sample distance of 0.25 mm and a spatial accuracy equivalent to 0.025 mm can be achieved [1]. According to [2], it is being applied and adopted in areas such as architecture, forensics, archaeology, and filming for the purpose of measurement. In addition, this technology has been implemented in the surveillance system to assist enforcers in the forensic investigation of crime scenes through the development of photogrammetry technology. Within the system, the camera or video becomes the tool to capture the crime scene of an incident. By using photogrammetry technology, the height of a suspect in a crime video footage taken from places such as convenience stores could be calculated scientifically [3]. This is supported by [4] who stated that photogrammetry can be used as digital evidence in court in terms of the crime scene measurements and photographic evidence in forensic analysis. Meanwhile, the infrared technology has been used in surveillance activities, particularly for night-time surveillance [5]. In this study, the most effective NIR filter for night time is identified so that it can be used in the camera of the surveillance system for the purpose of forensic mapping.

1.1 State-of-The-Art Surveillance System

A surveillance system includes a photographing lens system and a camera body to which the photographic lens system is detachably attached [6]. According to [7], the system uses physical observations, data collation, video footage, and many other monitoring instruments to carry out surveillance work. There are three existing types of system: (1) conventional system, (2) digital system, and (3) hybrid system. The conventional surveillance system usually used a videotape camera known as analogue camera. According to [8], analogue cameras turn the video signal into a format that can be received by a television or other receivers such as a Videocassette Recorder (VCR) or monitor. As for digital system, it is adequate with a digital video camera (without tape). High Definition (HD) format image and high resolution camera in megapixel also used IP-based camera, which digitizes the video signal using a specialized encoder that contains an on-board web server. The hybrid surveillance system is usually integrated with a biometric recognition system, infrared filter camera, thermal detection, and/or laser scanner. A hybrid system which integrates the NIR camera allows the system to function in two environments, day and night. Hence, a surveillance system that integrates night vision cameras usually uses NIR filter [9].

1.2 Brief of Infrared-Camera System

Infrared photography is defined as imaging by radiation in the near-infrared region of the spectrum. This technique works by recording scenes illuminated by invisible infrared radiation such as in surveillance, or by utilizing infrared radiation to reveal information that is invisible in the ordinary light such as the printing on a charred document [10]. Generally, the NIR cut-off filters are removed to increase the sensitivity under dark conditions. To remove the visible bands from NIR photograph, IR filters are mounted in front of the camera lens [11]. Filters placed in front of the camera lenses allow only certain wavelengths of energy to pass through the lens [12]. NIR cameras which are sensitive to near-infrared spectrum have many advantages. Therefore, in most NIR camera applications, spectral selectivity is required to maximize the amount of useful information. According to [5], the image contrast quality of night vision devices have greatly improved with the use of IR filters. When each filter is designed with a specific Electromagnetic (EM) wavelength, it allows only a particular range of EM wavelengths to pass through while other wavelengths are blocked. That is why each wavelength portion will have different reaction when exposed to different lighting.

1.3 Camera Configuration for Surveillance System

The camera set-up for a surveillance system depends on its application. There are several types of camera set-up used in previous researches that yield different results: (1) single camera, (2) stereo camera, and (3) multiple cameras. The single camera surveillance configuration is used to perform 2D imaging. In a research by [13], Pfleider utilized a 2D image analysis architecture with two complementary procedures for 2D tracking and initialization. Stereo Camera, on the other hand, requires two cameras in one baseline (set-up) to perform 3D-stereo imaging. One example is a research [14] which used a video stereo-image capture tool for photogrammetry. The “NuView Stereo Adapter,” a stereo-imaging device, and a high-definition video (HDV) were used in that study to capture stereoscopic video footage. The stereo adapter was mounted onto the video lens barrel and could be used to capture stereoscopic video footage in many photogrammetric applications, thus it becomes a complete stereo-imaging system. Next, the use of multiple cameras in a surveillance camera system provides a wide coverage area. Developed by a Massachusetts Institute of Technology (MIT) research group, MIT’s system uses a distributed set of sensors, and adaptive tracking to calibrate the distributed sensors, classify detected objects, learn common patterns of activity for different object classes, and detect unusual activities [15].
1.4 Subject Identification from Images

Based on [16], examples of forensic photogrammetry include determining the height of a bank robbery subject and the length of weapons used by the subject as depicted in the surveillance images. This is one of the identification techniques to identify criminals. In addition, subject identification of the suspect can be determined by applying biometric technique, anthropology (facial) measurement, and also thermal radiant of the image. According to [9], biometric recognition examines physiological or behavioural characteristics of a person to determine his or her identity.

1.5 Calibration of Surveillance Camera

In developing a high-accuracy surveillance camera system, a calibration technique must be implemented on the camera used. Based on the principles of photogrammetry, there are several techniques in camera calibration such as on-the-job calibration, self-calibration, and Plumb-line calibration [17]. On-the-job calibration is the term applied to a technique for determining the parameters of the lens and camera calibration in situ at the same time of the photography for the actual measurement of the object. The advantages of on-the-job calibration are most evident for non-metric cameras where there is a need to focus the lens for each epoch and, perhaps, for each exposure. Several researchers have applied this method in their studies such as [18, 14]. Meanwhile, self-calibration technique provides reliable values of the principal offset as well as the recovery of the focal length as the 90˚ roll of the camera offers improved configuration of the convergent images. This method is an extension of the concept embodied in on-the-job calibration. Hence, the observations of discrete targeted points on the object are used as the data required for both object point determination and for the determination of the parameters of camera calibration. This camera calibration technique is already well documented [19, 20]. Lastly, in Plumb-line calibration, the imaged distortions of an array of lines known to be straight in the object space are used to compute the parameters of radial and decentring lens distortion. However, this method is unable to determine the parameters of principal distance and principal point. This method is based on the premise that in the absence of distortion, a straight line in the object space will project as straight line in the image space.

2.0 EQUIPMENT AND SOFTWARE

The imaging system in this study consisted of three main components: a high-definition video (HD) camera, a set of near-infrared (NIR) filters, and an infrared (IR) illuminator. Figure 1 shows how the NIR filter was fastened onto the inside of the lens cone by a similar adapter ring and the infrared illuminator was wired to the video camera to provide NIR illumination when required. The system could operate remotely with or without any external power source.

![Figure 1 HDV attached with NIR filter](image)

2.1 HD Video Camera

Standard (720 x 560 pixels) and HD (1920 x 1080 pixels) video cameras can be used with the other equipment. However, HD’s capability was preferred because the large pixel count (2.6 x by 1.9 x) gave higher resolution for the same object distance. In this research, Sony HDR-HCSE with night-photography capability was used. Table 1 provides the imaging sensor specifications.

<table>
<thead>
<tr>
<th>Camera Type</th>
<th>Focal Length (mm)</th>
<th>Effective Format Size (mm)</th>
<th>Pixel Count (pixel)</th>
<th>Pixel Size (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sony HDR-HCSE</td>
<td>3.6-46.5</td>
<td>Still: 2.765 x 1.555 Video: 2.304 x 1.296</td>
<td>Still: 2304 x 1296 Video: 1920 x 1080</td>
<td>0.0012</td>
</tr>
</tbody>
</table>

2.2 Near-Infrared (NIR) Filter

Hoyair and X-Nile filters were used in the investigation. These devices block unwanted or scattered light from entering the lens of the camera, thus reducing image noise on the NIR video. Commercial night-time imaging illuminators mostly provide a wide NIR band ranging from 700 to 940 nm. An 800 nm filter removes the EM wavelengths up to 800 nm, so the bandwidth entering the lens is 800–940 nm. Tests have shown that an 850 nm NIR filter gives a good quality total-darkness scene.
2.3 NIR Illuminator

A few models of Sony video cameras have built-in NIR illuminators. However, external Sony NIR illuminators (Sony HVL-IRM Infrared Night Shot Light) were used in the research because they provide stronger and wider-angle (larger field of view) beams. In general, NIR illuminators are required in all night-time (low light or no light) tests.

2.4 Software

In this research, two off-the-shelf software were used for the processing. They are:

i. Australis – Camera geometric calibration software: Photometrix, VIC, Australia.

ii. MegaStat – Statistical analysis software (Microsoft Excel's add-in).

3.0 METHODOLOGY

This research was divided into two phases, which are phase 1 (geometric calibration) and phase 2 (case study).

3.1 Geometric Calibration

This phase included the determination of the principal point offset \( (x_p, y_p) \), lens focal length \( (c) \) or principal distance \( (PD) \), radial lens distortion parameters (usually \( k_1, k_2, \) and \( k_3 \)), lens decentring parameters \( (p_1 \) and \( p_2 ) \), and, in some instances, the dynamic fluctuation parameters [21]. In this study, the geometric calibration was divided into two: normal (colour) camera calibration (without filter) and NIR filter calibration.

3.1.1 Procedure

For the calibration, a high-precision calibration frame was constructed which consisted of 90 rods of height ranging from zero to 250 mm, equally spaced in a grid pattern over a flat piece of 600 mm x 600 mm x 20 mm square polycarbonate board (see Figure 2). Retro-targets of 4 mm of diameter were stuck onto the flat top of the rods. Two high-precision invar scale bars (800 mm and 220 mm) were placed on the board and each end of the bars had a retro-target of 6 mm of diameter.

Eight convergent (approximately 50–60° of tilt angle) photographs were captured for one set of data, taken from four positions which are each corner of the test field. Two photographs were captured at each position: landscape at 0° and rotation camera of 90° for portrait photo. The lens (focal length) was set as wide to maximize the field of view (FOV) of the calibration test field. The distance from the camera to the target plate was made sure to be less than 0.6 m (optimum distance). Hence, the photograph capturing was done in normal room lighting for normal calibration and low light-intensity environment for NIR calibration. The photograph acquisition was done in the same room. All the convergent photographs were then digitized (semi-automatic) using Australis photogrammetric software Version 6.06 package (Photometrix, VIC, Australia). As the study did not aim for high-precision 3D measurement in the order of few hundredths of a millimetre, the determination of lens distortion parameters was limited to \( K_1 \) (a residual radial lens distortion of approximately 0.4 pixel) and lens decentring parameters were discarded.

3.2 Case Study: 3D Measurement

The purpose of the case study was to implement the geometric calibration results. As the objective of the experiment was to identify the optimum NIR filter for surveillance application, the case study involved measurement on a mannequin and it was carried out according to the human body anthropometric concepts (e.g. height, facial, and body anthropometry) for the study on the accuracy of the 3D suspect identification data.
3.2.1 Procedure

Two different types of photographs, i.e. normal (colour) photograph and NIR photograph, with the same subject, which is a mannequin, were used in conducting this task. Initially, the mannequin (see Figure 3) was used to avoid any movement error during image capturing. The Sony HC5E HDV camera was used to capture convergent images of the mannequin. The camera is capable of taking colour and NIR photographs. It has a spatial resolution of 2 megapixels (effective still resolution) up to 6 megapixels (interpolated still resolution).

![3D Control Frame](image)

Figure 3 The mannequin was set up at a short distance in front of a control frame.

A built-in NIR light source which was located at about 10 mm below the lens was used to illuminate the target when the NIR photographs were taken. Four NIR filters with different wavelengths (715nm, 780nm, 830nm, and 850nm) were used. Sony HVL-IRC infrared illuminator was used to enhance the night vision. The focal length of the camera was set to a wide-angle setting which is 5.1 mm while for ISO, aperture, metering mode, exposure value, and colour tones, the settings were set to default and the shutter speed was set to auto mode. The camera was mounted on a tripod and its position was set 3.0 m away from the subject. Retro-targets were placed on the mannequin, which was positioned 0.1 m in front of the control frame. The control frame was used to provide an object-space control. The photograph of the mannequin was taken and the process was repeated three times for each NIR filter. The test was conducted in a low light-intensity environment. A semi-automated mono-digitizing technique using Australis software was used to conduct interior and exterior orientations of these convergent images. Test points of known coordinates were used to compare with the captured values to determine the 3D accuracy of the computed data. Basically, the tests were performed on the mannequin as the subject to determine the effect of colour and NIR photographs on the measurement results.

4.0 RESULTS AND ANALYSIS

The results of the evaluations are presented according to the order they were introduced in the methodology. The need to provide the 3D measurement accuracy is crucial because it is a significant factor in the development of a surveillance system, particularly for night-time applications.

4.1 Geometric Calibration: Results and Analysis

The final results obtained from the camera calibrations (normal and NIR) are tabulated in Table 2. The camera calibration results consist of four camera parameters, c (focal length or principle distance), x_p and y_p (principle point coordinates), and k_1 (lens distortion).

![Table 2 Camera calibration results](image)

<table>
<thead>
<tr>
<th></th>
<th>Mean from four observations (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>c</td>
</tr>
<tr>
<td>NIR 715 nm</td>
<td>5.1597</td>
</tr>
<tr>
<td>NIR 780 nm</td>
<td>5.1659</td>
</tr>
<tr>
<td>NIR 830 nm</td>
<td>5.1709</td>
</tr>
<tr>
<td>NIR 850 nm</td>
<td>5.1765</td>
</tr>
<tr>
<td>Normal</td>
<td>5.0916</td>
</tr>
</tbody>
</table>

All results were acquired by processing eight convergent images. Every measurement and processing stage used the same lens, object distance, focal length, and setting. The only difference was in the use of additional equipment such as different wavelengths of NIR filters. As a consequence, there were slight differences in the camera parameter values obtained from the observations. The analysis was carried out using ANOVA test for five sample variances (normal lens, 715, 780, 830, and 850 nm). Four camera parameters indicated by c, x_p, y_p, and k_1 were used as samples. To execute the hypothesis test, F-statistic results were used to determine whether the variances between the means of four samples are significantly different. The confidence level for the test was 95% (where α = 0.05). From the test, 95% confidence level showed that NIR filters with different wavelengths only
influenced \( c, \ x_p, \ y_p \), and \( k_1 \) parameters. Furthermore, most of the \( p \)-values gathered from the test were very small \( (p \leq \alpha) \). Hence, this provides a strong indication that different camera parameters \( (c, \ x_p, \ y_p, \ k_1) \) were obtained from the use of NIR filters with different wavelengths even though the differences were very small. Figure 4, 5, 6, and 7 show the differences in the camera parameters for every NIR filter.

The results from the five samples of camera calibration showed that most of the camera parameters for normal/colour imaging were shorter than that of NIR imaging due to the different wavelengths used. Since the object distance was set to be the same for both types of imaging, no error was expected from the object distance parameter.

### 4.2 Case Study: Results and Analysis

This stage involved the measurements obtained from the mannequin. Three-dimensional (3D) distances between the test points were used to compare the computed object space dimensional accuracy between NIR and colour photographs. The 3D distances were used instead of object-space coordinates to avoid any potential error that may originate from the slight displacement of the mannequin that was located in between the sets of convergent photography. The distances were regarded as true values and determined by a conventional method or direct measurement (Microscribe). Four longest 3D distances from the six test points were used for comparison (see Figure 8).

Each test consisted of four 3D distances \( (AD, \ CB, \ EF, \) and \( AB) \) and each 3D distance was measured using six types of measurement. Each measurement was repeated three times. Accordingly, seventy-two 3D distances were used in this test (see Table 3).
Table 3 Twenty-four 3D distances (mean from seventy-two 3D distances)

<table>
<thead>
<tr>
<th>Type of Measurement/Point</th>
<th>AD (mm)</th>
<th>CB (mm)</th>
<th>EF (mm)</th>
<th>AB (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microscribe (true value)</td>
<td>510.91</td>
<td>521.46</td>
<td>264.37</td>
<td>489.92</td>
</tr>
<tr>
<td>Normal Photography</td>
<td>510.20</td>
<td>520.80</td>
<td>263.84</td>
<td>489.52</td>
</tr>
<tr>
<td>NIR 715 nm</td>
<td>509.76</td>
<td>520.02</td>
<td>263.43</td>
<td>489.00</td>
</tr>
<tr>
<td>NIR 780 nm</td>
<td>511.99</td>
<td>520.22</td>
<td>263.51</td>
<td>489.18</td>
</tr>
<tr>
<td>NIR 830 nm</td>
<td>511.84</td>
<td>522.31</td>
<td>263.60</td>
<td>490.71</td>
</tr>
<tr>
<td>NIR 850 nm</td>
<td>510.14</td>
<td>520.74</td>
<td>263.93</td>
<td>489.58</td>
</tr>
</tbody>
</table>

The analysis of the test performed on the mannequin is presented in Table 4.

Table 4 The differences between true and measured values of 3D distances

<table>
<thead>
<tr>
<th>Set</th>
<th>Point A to D Difference ± (mm)</th>
<th>Point C to B Difference ± (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIR 715 nm</td>
<td>1.15</td>
<td>1.44</td>
</tr>
<tr>
<td>NIR 780 nm</td>
<td>1.08</td>
<td>1.24</td>
</tr>
<tr>
<td>NIR 830 nm</td>
<td>0.93</td>
<td>0.85</td>
</tr>
<tr>
<td>NIR 850 nm</td>
<td>0.77</td>
<td>0.72</td>
</tr>
<tr>
<td>Normal</td>
<td>0.71</td>
<td>0.66</td>
</tr>
<tr>
<td>Set</td>
<td>Point E to F Difference ± (mm)</td>
<td>Point A to B Difference ± (mm)</td>
</tr>
<tr>
<td>NIR 715 nm</td>
<td>0.94</td>
<td>0.92</td>
</tr>
<tr>
<td>NIR 780 nm</td>
<td>0.86</td>
<td>0.74</td>
</tr>
<tr>
<td>NIR 830 nm</td>
<td>0.77</td>
<td>0.79</td>
</tr>
<tr>
<td>NIR 850 nm</td>
<td>0.44</td>
<td>0.34</td>
</tr>
<tr>
<td>Normal</td>
<td>0.53</td>
<td>0.40</td>
</tr>
</tbody>
</table>

The 3D distance differences (mean) of the monodigitizing results were presented in relation to normal/NIR photography and measured points. Considering the whole dataset, most of the difference/mean accuracy values were lower than, or close to, ± 1.0 mm except for 715 nm and 780 nm NIR filters which demonstrated significant differences in the range of 1.08 mm to 1.44 mm for the measured points between AD and CB (see Figure 9). This phenomenon might be due to the low image quality for 715 nm and 780 nm NIR filter as compared to the others. Furthermore, each measurement of the 3D distance differences was totally different. The differences may have been caused by the differences in the camera parameters used by each NIR filters and colour photography.

The difference (mean) accuracy values of 850 nm NIR filter photographs were similar to the colour photographs (less than ± 0.8 mm). This suggests that 850 nm NIR filter photography has the potential to be used for acquiring 3D spatial data and taking measurements in low light intensity conditions. The standard deviation for all data was consistently distributed within the range of 0.1 to 1.0 mm (see Figure 10). The ANOVA test (α (0.05)) showed that the differences between the means were significant (H0: µ1=µ2 was rejected).
5.0 CONCLUSION

Based on the geometric calibration results and analysis, three remarkable findings were found in this study. First, most of the calibrated parameters were found to be slightly higher for NIR imaging due to the different wavelengths used. Secondly, the calibrations of four sets of NIR filters had produced a small variation in the camera parameters. Thirdly, 850 nm NIR filter was found to be the most promising and suitable filter for surveillance in low light conditions (or in darkness). The clarity of the images displayed was of the finest quality as compared to other NIR filters.

Based on the results and analysis of the case study, the 850 nm NIR filter gave better measurement accuracy compared to the other NIR filters. Physically, the images taken from NIR filter 850 nm were clearer, sharper, and smoother, whereas in the night surveillance system, light disturbance may have certain effects on the camera or video which consequently affected the visual image and resolution of the system.

6.0 RECOMMENDATIONS

Further studies may be carried out in the area of NIR emitter beam-width, beam-strength, and beam-bandwidth. An NIR emitter provides invisible NIR light similar to a flash light or spot light. Beam-strength is a major factor, in which an emitter set at a higher beam-strength often produces a more whitish image.

Acknowledgement

The authors would like to express their sincere appreciation to Universiti Teknologi Malaysia and the Ministry of Higher Education (MOHE) of Malaysia for giving the UTM research grant GUP Tier 1 (Vol 05H42) to conduct this research project.

References