The Comparison of Capabilities of Low Light Camera, Thermal Imaging Camera and Depth Map Camera for Night Time Surveillance Applications

Karol Jędrasiak and Aleksander Nawrat

Abstract. Night time surveillance is an important task for existing monitoring solutions. There is a need for low cost and high light-insensitivity solution. In this paper we present a results acquired during night indoor surveillance using three common vision devices and RGB-D camera. Experiments were performed in full light, minimal light and pulsating light simulating light alarm. Additionally 20 people were asked to recognize person in the selected frames of the acquired video sequences. Comparison of detection results and questionnaires answer charts are presented. It is presented that RGB-D cameras show great potential for low cost constant autonomous indoor surveillance regardless of the light conditions in the room.

Keywords: camera comparison, low-light camera, thermal imaging, RGB-D camera.

1 Introduction

Night time surveillance is an important task for existing monitoring solutions. Minimal light amount received by video cameras detectors and low signal to noise ratio (SNR) makes it difficult to detect and identify burglars. Existing solutions are usually based on visible light video cameras which are suitable for usage during day however they experience problems during twilight and are useless during night. Using additional artificial lights during night in order to improve vision conditions is the most basic solution, however it is a costly solution and not commonly used. Recently, low-light cameras are getting more popular. Low-light camera is described as a video camera capable of observing the scene characterized by a illumination

Karol Jędrasiak · Aleksander Nawrat
Silesian University of Technology, Institute of Automatic Control,
Akademicka 16, 44-101 Gliwice, Poland
e-mail: {Karol.Jedrasiak,Aleksander.Nawrat}@polsl.pl

A. Nawrat et al. (Eds.): Advanced Technologies for Intelligent Systems, SCI 440, pp. 117–128.
springerlink.com

lower than 1 lx. Night time capabilities of most of low-light cameras can be further improved by using IR (infrared) illuminators.

For surveillance of strategic buildings and infrastructure, thermal imaging cameras are used. Using thermal cameras is the most costly solution but at the same time the most reliable. Thermovision detectors are unaffected by visible light conditions, therefore they work exactly the same regardless of the day/night and weather conditions.

There is a need for low cost and high light-insensitivity solution. Recently more attention are getting RGB-D cameras, where D stands for depth map acquired from e.g. structural light analysis. Cheap IR projectors can be used to create a video camera suitable for usage in short range but unaffected by light conditions.

There is a need to compare the mentioned video cameras in terms of intruder detection quality during night. In this paper we present a results acquired during night indoor surveillance using four mentioned vision devices. As a quality measurement number of positively detected to all detected frames ratio was assumed. Additionally we decided to ask 20 people interested in computer vision whether they can recognize person in the acquired frames. Results of experiments and questionnaires are presented.

2 Night Time Surveillance Applications

Most of surveillance solutions are based on visible light cameras. It is difficult to perform traditional background subtraction methods for motion and object detection due to small size and low contrast of objects of interest. Using of contrast change instead of pixel intensity change for object detection was presented in [1]. False detections were suppressed using motion prediction and spatial nearest neighbor data association.

Alternative approach is to previously filter night time images in order to use traditional background subtraction methods [2]. A method exploiting the fact that most of surveillance systems continuously operative is presented in work [3]. Night time images are enhanced using day time luminance and reflectance. In paper [4] authors suggested use of independent component analysis (ICA) for indoor motion detection using background subtraction. First the background image is used to train ICA model, then the ICA is used to separate foreground and background images.

Some solutions use more than single camera in order to increase the quality of the detection. In paper [5] a stereo imaging system of low-light cameras and IR illuminators was used to detect and measure craniofacial anthropometric measurements with an accuracy of 27 mm within 3m range. Strong shadows and light reflections during night time conditions are disturbing motion detection. Shadow removal algorithm used in [6] is based on an assumption that equation (1) is fulfilled. The DELTA value was experimentally set as 0.1.

$$max = \left\{ \left| \frac{R_{fr}}{R_{bg}} - \frac{G_{fr}}{G_{bg}} \right|, \left| \frac{G_{fr}}{G_{bg}} - \frac{B_{fr}}{B_{bg}} \right|, \left| \frac{B_{fr}}{B_{bg}} - \frac{R_{fr}}{R_{bg}} \right| \right\} \leq \delta, \tag{1}$$

where:

R_{fr}, G_{fr}, B_{fr}, current frame values,

R_{bg}, G_{bg}, B_{bg}, background model values.

There is also a wide research area of using near-infrared cameras and thermal imaging cameras for night time surveillance. Both types of cameras can operate in conditions which human eye perceives as a total darkness. In paper [7] it is presented that traditional motion detection algorithms work well on illumination invariant images acquired from the mentioned video devices. Work described in [8] shows that simple calibrating thermal imaging camera for range of temperatures 30C-40C is sufficient for object detection using simple thresholding therefore easily done in real-time. Similar approach was used in home alone faint detection [9].

Thermal imaging cameras field of view (FOV) is relatively small (in range 10-20 degrees) in comparison to visible light cameras. To surpass this limitation in paper [10] omnidirectional system was presented. Wide FOV was acquired using hyperbolic IR mirror.

3 Vision Devices

Usually visual surveillance is based on video cameras. Images acquired using video cameras are two dimensional matrices composed of sensor responses that can be described by (2) as presented in [11]:

$$p^x = e^x \cdot n^x \int_{\omega} S^x(\lambda) E(\lambda) R(\lambda) \, d\lambda, \tag{2}$$

where:

p^x,3-vector of sensor responses,

S^x, surface reflectance,

E, spectral power distribution of the illumination,

R, 3-vector of sensitivity functions of the video camera,

ω,visible spectrum range

e^x,unit vector in direction of the light source,

n^x,unit vector corresponding to the surface normal at x,

Video device is therefore every device which output signal is also a two dimensional matrix of 3-vector responses.

The most popular vision device is a light video camera. For experiments we used a Logitech QuickCam camera (fig. 1). Its equipped with CCD sensor and its output format consists of one byte per R, G, B channel. Image resolution is 320x240 pixels. Video frame rate is 25 frames per second. Its parameters are typical for webcam cameras.

The low-light camera we used was a Sony IPELA SNC-RZ50N model (fig. 2). It is a PTZ camera with a maximum image resolution of 711x485 pixels and night-

Fig. 1 A Logitech QuickCam webcam

Fig. 2 A Sony IPELA SNC-RZ50N IP camera

mode function. Using night-mode it is possible to acquire clear images in low-light environments around 0.3 lx. Depending of the current zoom level camera's FOV is in range from 1.7 to 42 degrees.

The long-wave thermal imaging camera used was Thermoteknix 110K model (fig. 3). Its detector resolution is 284x288 and frame rate can be as high as 50 frames per second. Detector pitch is $25\mu m$ and spectral response is in range $8-12\mu m$. As a thermovision camera it is visible light invariant therefore it is best suited for continuous day/night time surveillance applications. Unfortunately its FOV is only 15.6 x 11.7 degrees.

RGB-D are video devices which output is depth map of the observed scene. At the time of research there were three main methods of acquiring depth maps. First and the most popular was to use stereo vision of two calibrated cameras. Second method is counting length of flight of emitted photons. This method is used by time-of-flight cameras. Recently a lot of attention received generating depth maps from structural light because of Microsoft Kinect device (fig. 4).

We have measured that Kinect is using near-infrared 830 nm (fig. 5) illuminator in order to display a known pattern onto the scene and a IR camera is used to calculate the depth. IR camera FOV is 57x43 degrees. It is believed [12] that depth is computed using triangulation against a known pattern from the IR projector. For each pixel in the acquired image a correlation window is used to compare the local

Fig. 3 A Thermoteknix 110K thermal imaging camera

Fig. 4 A Microsoft Kinect game controller (RGB-D camera)

pattern with the pattern at the known depth. It is used to compute disparity and the depth is computed using equation that comes from stereo vision systems (3):

$$z = \frac{b \cdot f}{d}, \tag{3}$$

where:

z, depth at the pixel point with disparity d,

b, baseline between stereo vision system cameras,

f, common focal length.

Depth data acquired from RGB-D camera can be used in fusion with RGB camera in order to perform real time segmentation and tracking of objects in 3D space rather than its 2D projection onto the camera lenses [13]. It also possible to build rich 3D indoor maps using RGB-D mappings [14].

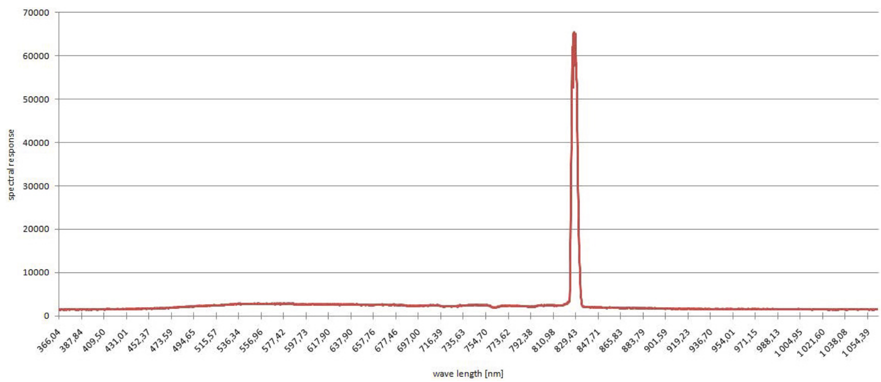

Fig. 5 Spectral response of the Kinect's IR projector

There are also other vision devices but they are not so well researched as the ones presented above. One of the most interesting are the 3D scanners based on radio frequencies echoes.

4 Experiments Description

Goal of the tests was to compare potential of tested devices for indoor surveillance during night. Experiments were divided into two parts. First one tested how easy information from the cameras can be used by standard motion detection algorithms for autonomous object detection. Second test's goal was to choose which camera is best in potential monitoring operator for night time surveillance.

4.1 Motion Detection by a Computer

During part one, three partial experiments were performed. For all three experiments the observed scene remained the same (fig.6). In the left side of the image there are multiple windows. On the right side, there is a door to the inner corridor of the building. In the background, various items and a wall can be seen. Actor's task was to jump in through the window in the left side of the scene then sneak around to the door in the right side of the room. All four video devices streams were recorded at the same time using developed video recording application.

As a quality factor a ratio r (4) computed as ratio of all frames with a person to detected person frames was used:

$$r = \frac{\sum_{i=0}^{n} I(n)}{n}, \tag{4}$$

Fig. 6 Scene used for the experiments

where:

n, number of all detected frames by motion detection algorithm,

$I(n)$, binary function that takes as input index of the frame and return 0 or 1 depending if the frame was tagged as a detected.

As a motion detection algorithm for visible and RGB-D cameras we used standard background subtraction algorithm with background image modeled as follows (5):

$$B_{t+1} = \alpha I_t + (1 - \alpha) B_t, \tag{5}$$

where: B_{t+1}, segment background value,

I, segment median intensity,

α, speed of background update.

For RGB-D cameras additionally empty black pixels were omitted during background updating.

However, as suggested in [6] for thermal camera, it was decided to use a different motion detection algorithm. The global image average as a reference value for thresholding. Each pixel with intensity value different then average background value by beta parameter was assumed as a foreground pixel. Beta parameter was experimentally set as 0.3.

First experiment was performed with the light on as a training run for the motion detection algorithms and the actor. Sample frames from the trial were used for questionnaire. Second experiment was performed during night with the lights off. In almost complete darkness actor's task was to jump in through the window, sneak through the whole scene and escape via the door in the right side of the observed scene. Finally during the third experiment the flashing light was used. It was an imitation of a visible light alarm that could be equipped in e.g. stores. During all trails actor performed the same task only the light conditions were different.

4.2 Motion Detection by a Human

Motion detection in indoor nighttime surveillance applications is usually performed by a trained human operator. In order to compare the results of automatic motion detection and a natural human ability to detect other humans in static images from different vision devices we created a questionnaire. Twenty potential monitoring operators were asked to answer simple questions whether they can see a person and if they can distinguish gender of the person. Images used in questionnaire are presented in fig. 7. Acquired results are presented and compared with the automatic results in the following section.

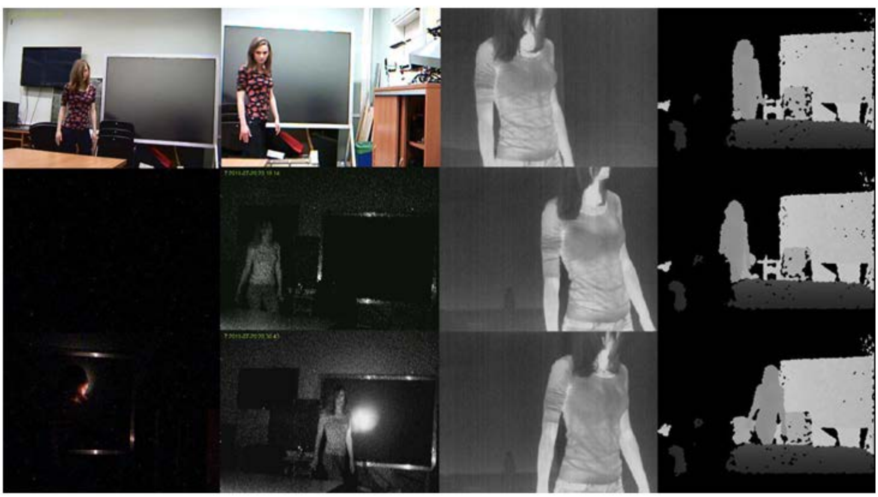

Fig. 7 Images from the four vision devices used during experiments that were chosen in order to compare automatic and manual results using questionnaire

5 Results

After the first experiment performed in the full light conditions it was no surprise that all video cameras scored high result. The visible light camera's score was 0.959. The ratio for low light camera was 0.957 and for thermovision 0.967. Bright scene, no shadows and camera movement are optimal conditions for background subtraction algorithm. RGB-D camera scored a result 0.961. It is a value comparable with traditional video devices. Results of the corresponding question from the questionnaire are presented in fig. 8. It can be seen that potential monitoring operators were able to easily fulfill the task of finding person in the images from the traditional video devices. RGB-D cameras produce a scene visualization that is different type than video devices, therefore respondents were slightly confused and their answers were less confident.

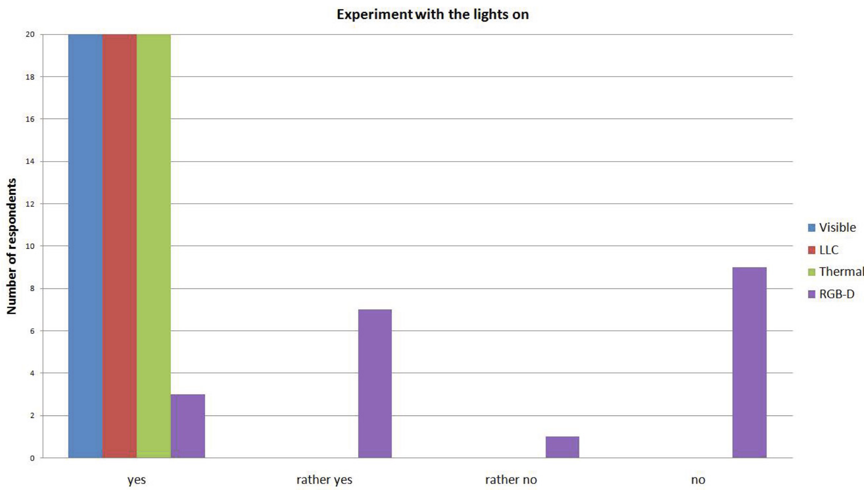

Fig. 8 Results of the questionnaire's part related to the first experiment

Experiment number two was performed in the conditions perceived by a human as almost complete darkness. Therefore images acquired from visible light camera contains only noise and result ratio is equal to 0. However in images from low light camera motion detection algorithm was easily able to detect person. The ratio was 0.903 which is comparable result with the full light experiments. There was no difference in images acquired using thermal imaging and RGB-D cameras. Human respondents also were not able to detect person in images from visible light camera. Detailed results are presented in fig. 9.

Experiment with the pulsating light imitating light alarms often installed in stores acquired interesting results. Human respondents' answers (fig. 10) were comparable with those of the previous experiment. However video devices results were affected. Visible light camera's results in darkness was the same. Pulsating light affects all pixels intensities that are visible in the images from low light cameras therefore successful detection occurs only in the interval between pulses. The result ratio was measured as 0.531 while thermovision and RGB-D cameras proved to be illumination invariant and therefore optimal for indoor surveillance. Their results were respectively 0.981 and 0.983.

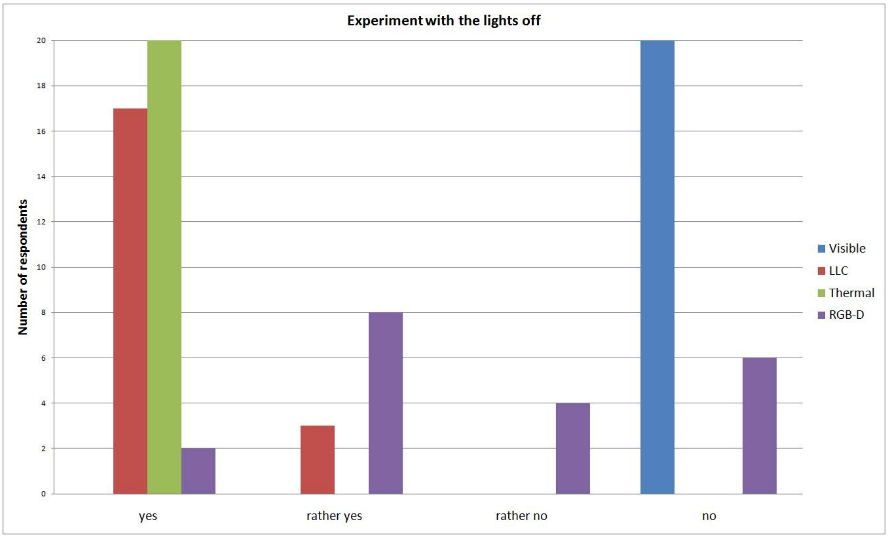

Fig. 9 Results of the questionnaire's part related to the second experiment

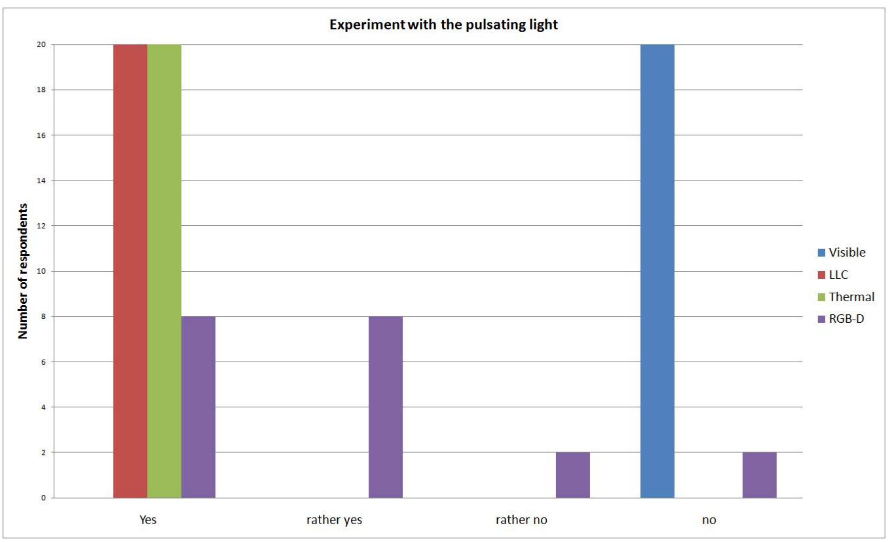

Fig. 10 Results of the questionnaire's part related to the third experiment

6 Conclusions

In this paper the three traditional surveillance video devices were tested against an alternative solution based on RGB-D camera. Capabilities for indoor day/night surveillance applications were tested during three experiments and their goal was to detect person in the acquired image sequences from the camera. Additionally 20 people were asked to answer simple questions about images acquired during the experiments. Human and traditional motion detection by background subtraction algorithms capabilities for object detection were compared and the results are presented. Acquired results shows that humans are not used to different type of scene visualization. Vision devices are compared using a quality ratio computed as a successfully detected frames to all possible for detection frames ratio. It is shown that both expensive thermovision camera and very low cost RGB-D camera are illumination invariant and ideal for indoor surveillance during day and night. It was also shown that existing solutions based on LLC are vulnerable for often installed close light alarms. Additional RGB-D camera feature is that experiments were based only on two dimensional scene projection. Acquired depth map makes it possible to use alternative motion detection algorithms in three dimensional space in order to significantly increase detection quality. Also anthropomorphical measurements could be performed for e.g. person identification.

In the future, RGB-D cameras capabilities will be tested in large halls and outdoor in order to precisely measure system limitations. Alternative way of research is developing motion detection and recognition algorithms that utilize three dimensional depth data acquired from low cost RGB-D cameras.

References

1. Huang, K., Wang, L., Tan, T., Maybank, S.: A real-time object detecting and tracking system for outdoor night surveillance. Pattern Recognition 41, 432–444 (2008)
2. Wang, Y., Fan, C.: Moving Object Detection for Night Surveillance. In: Proceedings of 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 236–239 (2010)
3. Yamasaki, A., Takauji, H., Kaneko, S., Kanade, T., Ohki, H.: Denighting: Enhancement of nighttime images for a surveillance camera. In: 19th International Conference on Pattern Recognition, Sapporo, pp. 1–4 (2008)
4. Tsai, D., Lai, S.: Independent Component Analysis-Based Background Subtraction for Indoor Surveillance. IEEE Transactions on Image Processing 18(1), 158–167 (2009)
5. Chong, A., Ariff, M., Majid, Z., Setan, J.: Night-time surveillance system for forensic 3D mapping. In: 2010 3rd International Congress on Image and Signal Processing, pp. 502–506 (2010)
6. Gupta, A., Ashis Pati, K., Venktesh, K.: Indoor Multi-Camera Surveillance System, Mid Term Project Report. Indian Institute of Technology, Kanpur (2009)
7. Ariff, M., Majid, Z., Chong, A.: Near-Infrared Camera For Night Surveillance Applications. Geoinformation Science Journal 10(1), 38–48 (2010)
8. Wong, W., Tan, P., Loo, C., Lim, W.: An Effective Surveillance System Using Thermal Camera. In: Proceedings of International Conference on Signal Acquisition and Processing, ICSAP 2009, Kuala Lumpur, Malaysia, pp. 13–17 (2009)

9. Wong, W., Lim, H., Loo, C., Lim, W.: Home Alone Faint Detection Surveillance System Using Thermal Camera. In: Proceedings of the 2010 Second International Conference on Computer Research and Development, Kuala Lumpur, pp. 747–751 (2010)
10. Wong, W., Tan, P., Loo, C., Lim, W.: Omnidirectional Surveillance System Using Thermal Camera. Journal of Computer Science and Engineering 3(2), 42–51 (2010)
11. Schaefer, G.: How useful are colour invariants for image retrieval? Computer Vision and Graphics, Computational Imaging and Vision 32, 381–386 (2006)
12. Internet, `http://www.ros.org` (visited June 1, 2011)
13. Bleiweiss, A., Werman, M.: Fusing Time-of-Flight Depth and Color for Real-Time Segmentation and Tracking. In: Kolb, A., Koch, R. (eds.) Dyn3D 2009. LNCS, vol. 5742, pp. 58–69. Springer, Heidelberg (2009)
14. Henry, P., Krainin, M., Herbst, E., Ren, X., Fox, D.: RGB-D Mapping: Using Depth Cameras for Dense 3D Modeling of Indoor Environments. In: Proceedings of the 12th International Symposium on Experimental Robotics, Delhi, India (2010)