Automatic Segmented Area Structured Lighting

Kruti Goyal, Hadi Baghsiahi & David R Selviah

Department of Electronic and Electrical Engineering
University College London (UCL), Torrington Place WC1E 7JE London
d.selviah@ucl.ac.uk, http://www.ee.ucl.ac.uk/staff/academic/dselviah

Abstract

The aim of the research is to devise an automatic way to view and segment a scene of discrete 3D objects with or without ambient illumination and then to fully illuminate each object in turn without illuminating other objects or the background. The structured illumination must be controlled in time and space to have the same shape, size and position as the object. There is a need for such a system in the entertainment and visual arts industries for sound and light shows at night outdoors for selectively illuminating buildings, in caves or caverns for illuminating rock formations, or for illuminating mannequins, statues or waxwork figures in theatres sequentially in synchrony with a voiceover narration discussing each in turn. In these applications, such a technique has an advantage over the use of spotlights as only the object of interest is illuminated and not nearby objects or the background so helping the viewer to concentrate on just the object of interest. In this paper a video camera and projector system is reported with real time image processing feedback via a computer. The way in which the image processing algorithms in the feedback loop were developed to overcome various issues is explained.

Keywords: Object-Background segmentation, Structured illumination, Spotlight, Signal processing, Image alignment

1 Introduction

The research aims to design and demonstrate a camera and projector intelligent system to illuminate 3D Objects within a 3D scene while minimizing illumination of other nearby objects and the background of the scene. The system should be capable of illuminating each individual object alone and alternatively illuminating each individual object in a sequence in turn. The system can be used to illuminate all of the objects in the scene at the same time or a single object can be illuminated while the rest of the objects and the background are under background are kept dark. The research has various applications in the entertainment industry such as in outdoor sound and light shows [1] in which individual buildings [3] or statues are illuminated in turn either in time to music or synchronised with explanations about the objects given by a speaker [2], in underground cave tours where individual rock formations are illuminated in turn as a speaker explains them [4], recently it has also been popular to project images onto mannequins and statues

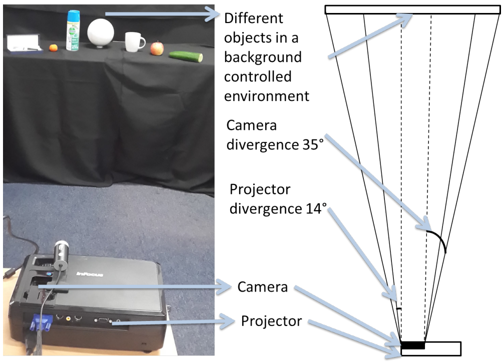

Figure 1: Experimental Arrangement and Bird view of the arrangement

to make them appear to be talking. In such situations, the events usually take place at night or in dark rooms and the difficulty is how to arrange the projector to mainly illuminate just one object without illuminating other nearby objects. The alignment and calibration of the projector can be very time consuming. In this paper we describe the development of a camera and projector system with image processing software to carry out this automatic alignment of the projected image and the 3D objects. In this paper only uniform illumination is projected although in future it could be projected moving images.

The arrangement used for the paper is as shown in figure 1. A video camera is placed on top of the projector, ans is secured to it, and the output of the camera and input of the projector are connected to a HP Pavilion DV6 3107ax laptop computer which requires 90 W AC power and has 500 GB memory. The video camera was chosen to be Microsoft Lifecam Cinema which has a USB 2.0 connection, a speed of 30 fps (frames per second) and is powered by the computer through the USB 2.0 link. The resolution of the camera was 960 × 544. Thw projector chosen for the project was an Infocus X9 DLP project with a brightness of 1800 ANSI lumen and 1400 × 1050 resolution. The camera is rotated so that its horizontal axis matches that of the projector so that the captured images do not need to be rotated. The camera is placed on the top of the projector to minimise the distance between the two. In the experiments reported in this paper the distance between the center of the lenses of the camera and projector is 8 cm. This distance is measured vertically as the camera lens was placed directly above the projector lens. A dark background is used to simulate the typical environments used in sound a light shows. This also helps to achieve proper object background segmentation due to the intensity difference between the background and the objects to be illuminated. A light background and dark foreground can be used alternatively, in which case the picture needs to be inverted before further processing.

The paper focuses on three main issues which need to be solved. Firstly, the system should be able to automatically detect and separate the objects from their background. Secondly, the projector must project image such that it matches the objects detected by the camera exactly. Finally, only the specified objects should be illuminated at a particular time keeping all of the other objects and the background in the dark. This whole process should be carried out automatically in realtime and should be robust to changes in the ambient light. In order to achieve this the camera-projector system is used in conjunction with image processing algorithms written in LabVIEW is used.

The surroundings contain 3D objects, therefore, the detection technique should be such that the appropriate information of the surroundings is retrieved when the 3D environment is converted into a 2D image. Laboratory Virtual Instrument Engineering Workbench (LabVIEW) is graphical programming language used for computer vision and image processing and was chosen over MatLab due to its speed and graphical user interface which enables various functions such as different filters to be used without changing the code or need to recompile the code and due to LabVIEW's better hardware-software interface. LabVIEW image acquisition was used to capture the surroundings and the image was then processed as presented in section 2. The next step is to determine the best match between the projected illumination and the object such that the projected light is exactly on top of the object. Normalised cross correlation between the transformed captured image with and without projection was used for this purpose. The disadvantages of cross correlation includes sensitivity to illumination, rotation and scale changes. However, in this research they prove to be advantageous as due to the difference in normalised cross correlation value for different scale, illumination or rotation, the best match can be found. The objects were then selected and lit sequentially using Magic Wand Processing, a function provided by National Instruments as part of NI Vision, as presented in section 3.

The whole process takes about 35 seconds to one minute 15 seconds with a 4 GB AMD Phenom Quad-core processor and is independent of the number of objects used. However, the speed of processing is highly dependent on the processor used. Onboard FPGA image processing can give much higher processing speed than the laptop computer. The speed of the program also depends on the RAM available with higher RAM increasing the speed of processing significantly. The processing time is independent of number of objects detected, and, hence, any number of objects can be detected by this method.

2 Object-Background Segmentation and Image Processing

Each object needs to be segmented and separated from the background and other objects. Also the noise (unwanted particles and bright spots due to details of the scene and light sources and unwanted reflections) needs to be minimised [5–8]. LabVIEW 14.0 student edition together with IMAQ 2014, IMAQdx 2014, NI VISA 2014 and NI Vision acquisition 2014 and NI Vision assistant 2014 [9, 10], all provided by National Instruments, were used for this purpose as they conveniently have a variety of pre-written functions relating to image processing and computer vision and they provide various drivers for hardware integration.

2.1 Spatial Image Filtering

The captured RGB image is converted to greyscale and is segmentation filtered to separate the object from the background and from other objects. Different filters were tried to find the most appropriate filter. A Gaussian filter [11–14] was first applied to separate the objects from their background. It has the effect of keeping the important patterns while removing finer details, however, it also increases the noise present (figure 3). If the light is equally distributed across the image and bright, the Gaussian filter saturates and the whole image becomes white. In this case, the alternative is to use a low pass filter. The effect of the low pass filter is minimal, except that it smooths the boundaries of objects. If lowpass filter is used for figure 2, which is not uniformly lit, the object to background intensity difference will be low and the objects would not be detected, this is evident in later stages of image processing, that is after thresholding as shown in figure 4. Even when the kernel's (Convolution matrix determining the weight of the pixels under consideration and their neighbouring pixels) size is increased, the result is still the same as shown in figure 4. The final result after image transformations as described is section 2.2 results in a black image if a low pass filter is applied for non-uniformly lit environment. If no filter is applied, the result is the same as shown in figure 4. Hence, a Gaussian filter is suitable for non-uniformly lit enviorment such as that of figure 2, whereas lowpass filter is ideal for uniformly lit enviorment. Therefore, either a Gaussian or low pass filter is selected for object background segmentation depending on the ambient illumination conditions. A Gaussian kernel of size 3×3 with the following dimensions were used as they were found to give the best results.

$$G = \begin{vmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{vmatrix}$$

Figure 2: Captured image of the arrangement of 3D objects within the scene from the CCD camera

Figure 3: Image after application of a Gaussian Filter

Figure 4: Image after application of a Low Pass filter

2.2 Image Transformation and Morphology

The greyscale image is eroded [15] for four iterations to remove the finer details and bright spots, so that they do not connect two adjacent objects and make them appear as one, and also so that they are not mistaken as objects themselves. The number of iterations was varied and four was found to yield the best results. If a pixel under consideration has the binary value 1 and if all of its cardinal neighbouring (pixels to the immediate left and right, and top and bottom of the pixel under consideration) pixel values are 1 then the new pixel value is set to 1, otherwise it is set to 0. The erode operation is defined by:

$$X_{new} = \begin{cases} 0 \; if \quad x_{old} = 0 \\ 1 \; if \; x_{old} \; \& \; (Cardinal = 1) \end{cases}$$

Next, the image is dilated for four iterations, to bring it back to its original size. The number of iterations was varied and four was found to yield the best results. The dilation operation sets the new value of the pixel to be 1 if the original pixel is 1 or if its cardinal neighbours are 1, otherwise the value is set to 0. The dilate operation is defined by:

$$X_{new} = \begin{cases} 1 \; if \; x_{old} \; OR(Cardinal = 1) \\ 0 \; if \quad otherwise \end{cases}$$

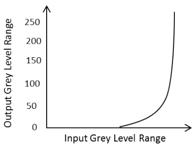

Figure 5: Exponential operation

A non linear exponential operation is applied to the pixel grey level intensities next using a look up table [16] to improve the contrast ratio and to remove noise due to intensity variation. This operation decreases the contrast and brightness in dark regions, whereas it increases the contrast in bright regions as can be seen from the non-linear curve in figure 5.
Next, the image is thresholded to separate the bright object from the dark background. A high and low threshold value is chosen, the pixels within this range is detected as the object whereas all the other pixels are darkened. As a black background is used and objects nearer to the camera are brighter than those further from the video camera, this essentially separates the object from its background. This step however, is greatly affected by the brightness of the room and the camera used, and needs to be varied in different ambient illumination. If the room is bright, the objects and background will reflect more light, and, hence, a high lower threshold value needs to be set. Whereas, if the room is dark, the reflected light from the objects and the background will be weaker, and, hence, a low lower threshold value needs to be set. Therefore, to avoid setting up a manual threshold every time the program runs, the maximum intensity of the greyscale image is found and a percentage of this (20 and 35 % were tried for the experiment) is taken as the lower threshold, the upper threshold is set to a constant, which is its highest value (255 for 8-bit image). Bright spots on the border of the image can be mistaken as objects and, hence, all the objects connected to the border are removed next. This step is optional, as in many cases the object can be near the border. The image is then equalised by distributing the pixels in intensity evenly and converting it to an $8-$bit image from a binary image. These image transformation operations are as shown in figure 6. These transformation were applied on images taken in a room with ceiling light and windows opposite to the scene. If the room is completely dark in the detection stage, the program would not be able to segment the objects from its background. However, the sequential illumination of the objects can be done in a dark room using the information obtained in the previous stages. The thresholding operation is given by:

$$i = \begin{cases} 0 \; if \quad p < t_{low}, p > t_{high} \\ 1 \; if \quad t_{low} \le p \le t_{high} \end{cases}$$

3 Calibration and Projection

The video camera has a greater field of view than that of the projector. To detect the area of projection in the captured image, a white image with the same pixel area as the projector is projected to illuminate the field of

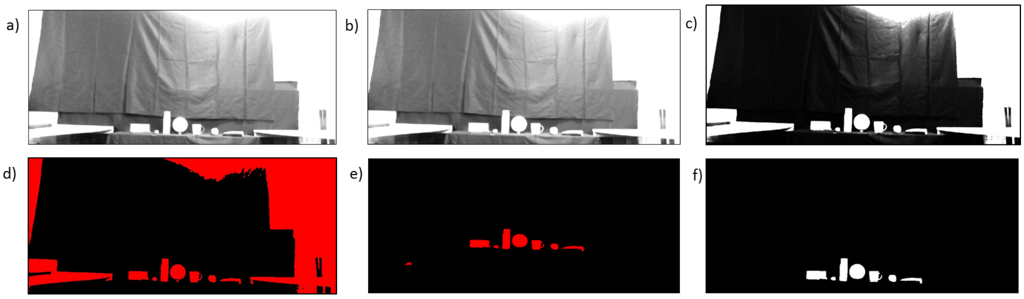

Figure 6: Image transformation: a) Erode b) Dilate c) Exponential d) Threshold e) Remove border objects f) Equalise

view of the projector. This area also called, region of interest (ROI), is then extracted from the captured image and image processing is subsequently only carried out within this area. This area selection can be carried out automatically by setting a high value for the lower threshold (around 230) to detect the area of projection, as it will very bright due to projection of a white image onto the scene. This area is then selected by using the magic wand function, which compares the intensity of the pixel under consideration with its neighbours and their neighbours, and selects all connected pixels which have a similar intensity within the set tolerance limit (20%). The coordinates for a particle under consideration are found using the particle analyser function in LabVIEW and correspond to the centroid of the bright particle with the maximum area. The maximum area is selected because there might be some other regions with high intensity which can be taken as the ROI resulting in the wrong selection. However, depending on the surface of the background, some regions in the projected area might not fall into the threshold range and are ignored, resulting in an error; this error, however, does not affect the final result significantly as the ROI selected is defined to be a rectangle.

After the ROI is extracted, one object is selected to be illuminated. Next, the processed image of the object is projected back through the projector onto the object. Automatic thresholding on the image can now be used (entropy [17–19] in this case), as the brightness of projected structured light is much greater than the other ambient illumination, hence, the ambient illumination can be ignored.

There may still be some difference between the edge of the area of the projected light and that of the real object due to the skew of the projector lens, the distance between the projector lens and the camera lens and error in locating and extracting the ROI. This error is corrected using a feedback loop. The scene onto which processed images are projected is continuously monitored using the video camera, and the latest captured and processed image (which is dominated by the illumination area shape) is correlated with the initial captured and processed image (Template). The dimensions and position of the projected image are altered until a maximum correlation is achieved. The use of one object instead of all, decreases the processing time and reduces errors. The projected structured light is rescaled, moved up and down and left and right, and the normalised cross correlation is calculated for each position and size. The final algorithm which moves the image up, down, left, right and varies the scale works well but it would be useful to incorporate additional transformations such as keystone to correct for the skew due to the 8 cm offset between the camera and the projector lens. Once the maximum correlation is reached when the projected image is exactly aligned with the real object, the calibration process ends. The dimensions and coordinates of the projected image are then recorded for further use.

The acquired illuminated image is then analysed using NI Vision particle analyser. Particle analyser can detect continuous regions or grouping of pixels with similar intensity, known as particles. The particle analyser then makes measurements of these regions or particles and also determine their coordinates. This is often referred to as BLOB (Binary Large Object) analysis. The particle analyser detects each particle and returns the coordinates of each object detected. The program then takes these pixel coordinates and uses magic wand processing which uses the intensity of neighbouring pixels, to select only one object and mask all the others, so that only one object can be illuminated at one time. The projected individual object images can then be projected sequentially to illuminate each object one after the other as required.

4 Experimental Results Discussion

The research investigated different algorithms, to select the best possible method. It was also conducted at different times of day to check the performance under different ambient background lighting conditions and at different locations with different 3D scenes to verify its robustness. The results obtained are discussed in this section. The area of interest is as shown in figure 7. The arrangement consist of a pen case, an orange, a bottle, a sphere, a mug, an apple and a cucumber, from left to right, the camera-projector system is 1.26 m from the objects. The sphere is illuminated for calibration purposes in figure 8 and 9, however, any object can be selected for this purpose. When the projection is of a different size than the object it is resized using a scale transformation as shown in figure 8. Next, the projection is moved in the x and the y directions to meet the edges of the object as shown in figure 9.

The objects were lit sequentially and all at once as shown in figures 10, 11, 12 and 13. The algorithm was also applied to different types of object as can be seen in the figures. The objects had various shapes, colours, textures and reflectivities. All of the objects were placed side by side at different positions on the table and could all be lit at the same time. To check the validity of the algorithm objects were detected and illuminated from different positions. It was found that the closest position the camera can detect the object and area of interest without becoming saturated is 14 cm. The longest distance, however, is only limited by the power of the projected light and in this case is about 15-16 m. There is no limit on the area of the largest detectable object, as long as the camera can detect it and it comes within the projector's field of view. The smallest object which can be detected without being eliminated as background noise is about 5 pixels or 0.14 pixels per degree according to camera's resolution, and, hence, very small objects can be detected if they are close to the camera-projector system but the same object is neglected as noise when it is far from the camera-projector system. Objects were placed close and far from the system simultaneously and illuminated as shown in figure 14 and the system was shown to illuminate them all correctly. Complicated and large shapes such as human beings can also be detected by the system as can be seen in figure 15.

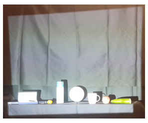

Figure 7: Projected image to detect ROI

Figure 8: Resizing of the projection to match the sphere

Figure 9: Shifting the projected light to the left

Figure 10: Light projected onto a Pen Case

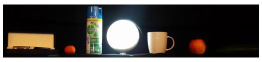

Figure 11: Light projected onto the Sphere

Figure 12: Light projected onto a mug

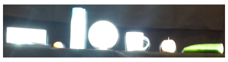

Figure 13: Light Projected on all the objects

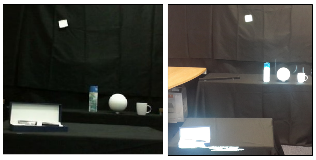

Figure 14: Arrangement (left) of and light projection (right) onto Pen case (closer to projector), post-it, bottle, sphere and cup

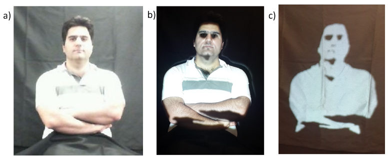

Figure 15: Detecting a complex subject: a) Scene with no projection on the person b) Light projected onto the person c) The detected person projected onto the background

The camera has a divergence of 35 degrees or horizontal divergence of 27 pixels/degree and vertical divergence of 16 pixels/degree, whereas, the projector has a horizontal divergence of 14 degrees or 100 pixels/degree and vertical divergence of 12 degrees or 87 pixels/degree. As the field of view of the camera is larger and the resolution is lower than the projector which decreases even more when part of the image is extracted, there is a difference between the processed image and the image which is projected the and this sometimes results in spillage of light into the background of about 1 pixel or more, even after the calibration is done. This is evident from the background illumination in figure 12 around the mug. This spillage cannot be seen by the camera due to the difference in divergence and point of view of the camera and can be solved by using another camera, which can view this spillage from another angle and counter it by manipulating the light to match the object.

5 Conclusions

The project's aim to make a robust system to illuminate discrete objects individually was achieved. The shifting and resizing process can be calibrated to enable the illumination to automatically illuminate an individual object to one pixel accuracy. The illumination became distorted when it is passes into the camera due to the magnification factor between the scene and the camera's detector array which affects the result of the filters and the result of correlation. This in turn affects the output shift and size of the projected image as due to distortion the position and size of the projection is deemed to be correct by the program, when in reality it may not be. These effects can be minimised by using a high resolution, anti-glare camera.

The camera used for this project has a resolution of 960 × 544 whereas the projector has a resolution of 1400 × 1050. In addition the angular divergence of the camera, 35 degrees, was much larger than that of the projector, 14 degrees, which aggravated the difference between their resolutions resulting in an angular resolution of 27 pixels per degree for the camera and 100 pixels per degree for the projector. Ideally the resolution of the camera in pixels per degree should be higher than that of the projector in pixels per degree. This can be achieved by a combination of selecting the number of pixels in the camera sensor plane and controlling the angular divergence by narrowing the cameras field of view using a zoom lens. Another limitation of this method is that some light was projected onto the background, which cannot be seen by the camera but is seen by the viewer at a different angle from the camera. The program's speed is processor dependent and takes about 35 seconds to 1 minute 15 seconds to complete with the current system. Therefore, the research was successfully able to achieve object-background segmentation, calibration and sequential projection. However, the aforementioned limitations such as projection on the backdrop of the object and effects due to environment; limit the application of this research and makes it viable for further investigation.

Acknowledgements

We express our gratitude to Dr. Sally Day for her support and her generosity in providing a work space and apparatus required.

References

[1] Dream makers. 2010. [Online]. Available from: http://www.soundandlight.com.eg/Shows/ GizaPyramidsHome/VideoPics/VideoTube2.aspx. [Accessed 20 February 2015].

[2] D. Bowies. 2015. [Online]. Available from: http://flavorwire.com/422969/deconstructing-david-bowies-diy-video-for-love-is-lost. [Accessed 20 February 2015].

[3] Shader Lamps. 2015. [Online]. http://web.media.mit.edu/ raskar/Shaderlamps/OtherImages/index.html. [Accessed 20 February 2015].

[4] Poole's Cavern. 2015. [Online]. Available from: http://www.poolescavern.co.uk/visitor-information/. [Accessed 20 February 2015].

[5] Szeliski, R. (2010). Computer Vision : Algorithms and Applications. Springer Science and Business Media, 1, 2, 3. London.

[6] Klinger, T. (2003). Image Processing with LabVIEW and IMAQ Vision. Prentice Hall Professional. Upper Saddle River, New Jersey.

[7] Pratt, W. (1991). Digital Image Processing, Second Edition. John Wiley and Sons, Inc. New York, NY. ISBN $0-471-85766$. pp $561-563$.

[8] Lewis, R. (1990). Practical Image Processing. Ellis Horwood Limited. New York, NY. ISBN $0-13-6383525$. pp $90-91$.

[9] Fairweather, I., Brumfield, A. (2011). LabVIEW: A Developer's Guide to Real World Integration. Chapman and Hall/CRC. Boca Raton, FL. ISBN 9781439839812.

[10] Posada-Gomez, R., Sandoval-Gonzalez, O.O, Sibaja, A.M., Portillo-Rodriguez, O., Alor-Hernandez, G. (2011). Digital Image Processing Using LabVIEW, Practical Applications and Solutions Using LabVIEW Software. Dr. Silviu Folea (Ed.). ISBN: 978-953-307-650-8. InTech, DOI: 10.5772/23285. Available from: http://www.intechopen.com/books/practical-applications-and-solutions-using-LabVIEW-software/digital-image-processing-using-LabVIEW. [Accessed 20 November 2014]

[11] National Instruments. (2000). IMAQ Vision for LabVIEW. Available from: http://www.ni.com/pdf/manuals/371007a.pdf. [Accessed 20 November 2014]

[12] Relf, C. G. (2004). Image Acquisition and Processing with LabVIEW. CRC Press LLC. Boca Raton, FL. 2^{ed}. USA. ISBN 0-8493-1480-1.

[13] Version, I., Bhuvaneswari, P., Therese, A. B. (2014). Edge Detection Techniques in Digital and Optical Image Processing. Int. Journal of Engineering Research and Applications. ISSN : 2248-9622. Vol. 4, Issue 5 (Version 3). pp.33-37.

[14] Ravi Kumar V. A., Nataraj, K. R., Rekha, K. R. (2012). Morphological Real Time Video Edge Detection in LabVIEW. (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (2). pp 3808-3811.

[15] Xiaobo, M., Jing, Y. (2011). Research on object-background segmentation of color image based on LabVIEW. 2011 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems, DOI: 10.1109/CYBER.2011.6011791. pp 190-194.

[16] National Instruments. (2000). IMAQ Vision Concepts Manual, (322916). [online]. Available from: http://www.ni.com/pdf/manuals/322916a.pdf. [Accessed 26 November 2014].

[17] Gonzalez, R.C., Woods R.E, S.L. Eddins. (2009). Digital Image Processing Using MATLAB. 2nd ed., Gatesmark Publishing. Knoxville, TN. Prentice Hall. Chapter 11.

[18] Loomis, J. (1998). (2015). Entropy. [online]. Available from: http://www.johnloomis.org/ece563/notes/basics/entropy/entropy.html. [Accessed 25 November 2014].

[19] Iliev, P., Tzvetkov, P., Petrov, G. (2014). Motion Detection Using 3D Image Histogram Sequences Analysis. [online]. Available from: http://nbu.bg/PUBLIC/IMAGES/File/departamenti/telekomunikacii/1.pdf. [Accessed 25 November 2014].