16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

IMAGE CHANGE DETECTION FOR A PERSONAL RAPID TRANSIT APPLICATION

Alan Peters, Torquil Ross-Martin†, Alin M. Achim**

*Department of Electrical & Electronic Engineering
University of Bristol, UK
†Advanced Transport Systems Ltd, Ashville Park, Thornbury, Bristol, UK

ABSTRACT

Automatically identifying objects and people left in the interior of vehicles is highly desirable because human monitoring has high running costs and low efficiency associated with it. A new Personal Rapid Transit (PRT) system currently being designed by Advanced Transport Systems Ltd (ATS) features many autonomous vehicles and therefore the task is of particular importance. This paper describes two approaches that use changes in the visual image of the interior to predict the likelihood of left objects and remaining people. The first approach is based on identifying structural differences. The second approach uses a shading model method. A variation of the shading model with information from the colour channels is also described. The results show that the modified shading model approach gives the best performance.

1. INTRODUCTION

Whilst the task of identifying objects in the path of autonomous vehicles has been heavily investigated, less research has been conducted on the task of identifying objects or persons left in the interior of a vehicle. This paper presents two methods using image change detection techniques.

Automatic occupancy detection is highly desirable in Personal Rapid Transit (PRT) systems. The Urban Light Transport system (ULTra), designed by Advanced Transport Systems Ltd (ATS) is a new PRT system that features autonomous vehicles carrying up to four people. When a vehicle reaches its destination, it is necessary to determine that all passengers have disembarked. It is also necessary to know that there are no items left in the interior, both for the convenience of the passengers and also for security reasons. It is desirable that this detection system is automatic so that a human operator is not required to make decisions every time a vehicle leaves a station. Suggested approaches include weight or heat measurement but in this paper we restrict ourselves to two approaches both based on finding the difference mask between two image frames captured by onboard cameras.

ULTra uses driverless battery-powered vehicles operating on a network of concrete guideways [1]. Passengers arrive at a station and programme a waiting vehicle to take them automatically and non-stop to a desired destination. ATS is currently delivering a small-scale system to Heathrow Airport Terminal Five. The ULTra development programme has included investigations into guidance and obstacle detection systems. Much of this work was done as part of the Autotaxi programme [2, 3].

The proposed detection system uses two cameras mounted on the underside of the vehicle's roof. The images from these cameras are then analysed to determine the differences between how the vehicle interior looks and how it should look when it is empty. The cameras are required for surveillance so it is efficient to also use them for occupancy and object detection. The use of image analysis techniques is further supported by the continued development of more powerful computer processors.

Recent surveys on image change detection were conducted by Radke et al. [4] and Dai and Khorram [5]. Previous surveys were written by Coppin and Bauer [6] and Singh [7]. The latter two concentrate on remote sensing in forest monitoring applications.

The rest of this paper is organised as follows. In Section 2, the problem is more formally defined. In Section 3, the methods are described and in Section 4, results are presented. Section 5 presents conclusions.

2. PROBLEM STATEMENT

2.1 Image Change Detection

The aim of the proposed algorithm is to identify the change mask between an image of an empty vehicle and an image of the vehicle before it leaves the station. Any changes due to people or objects should be treated as significant. Any changes due to illumination changes should be ignored and treated as insignificant.

Let us denote the empty (template) image I_1 and the test image (image obtained by the cameras) I_2. Both images map a pixel coordinate x to an intensity or colour. The binary change mask at pixel x, B(x), can be defined as follows

$$\mathrm{B(x)} = \begin{cases} 1 & \text{if significant difference between } \mathrm{I_1(x)} \text{ and } \mathrm{I_2(x)} \\ 0 & \text{otherwise} \end{cases}$$

The change mask can then be used to determine the presence or otherwise of persons or objects in the vehicle.

In this application the template image will be updated when the controller knows the vehicle to be empty. This is likely to be when the vehicle is initialised at the beginning of an operating session. The algorithm will use the data from the two cameras seperately because the field of views do not overlap (the two cameras cover two different regions of the interior). The vehicle is defined as empty if the algorithm decides that both ends of the vehicle are unoccupied.

2.2 Key problems

The system must decide on the status of the vehicle in a minimal time (less than 0.5 seconds is desirable). It must not use significant memory or processing resources of the vehicle onboard computer. Ideally the system should make any decision using only the template and test images rather than a series of frames from a video sequence.

Figure 1: Interior of PRT vehicle with marked area of interest. The shadows caused by strong sunlight can be easily seen.

The system should also work in a variety of changing environments. This includes changing light levels and weather conditions. Any resulting illumination changes should be treated as insignificant.

3. CHANGE DETECTION ALGORITHMS

Two methods are implemented and analysed. The first is based on the structural similarity measure [8] which was devised as a way of measuring the perceptual similarity between two images. The second is based on the shading model method which is a method for detecting change developed by Skifstad and Jain [9].

3.1 Pre-processing

Pre-processing is necessary to reduce the effects of illumination changes. First a mask is applied to remove the window and door sections of the interior (cf. Figure 1). This is the same size for all vehicles. These areas will change depending on the outside environment and therefore the change is insignificant.

A low-pass smoothing filter is applied to reduce the effect of irregular pixels and aligns the images if they are misaligned by sub-pixel amounts. In initial experimentation this filter is a spatial averaging filter with an averaging mask size of 10 x 10 pixels. Homomorphic filtering was experimented with as an attempt to reduce low frequency illumination changes but little performance gain was observed.

3.2 Absolute differencing

The simplest method for change detection uses the absolute difference between corresponding pixels. The difference image and the binary change mask are as follows

$$\mathrm{D(x)} = |\mathrm{I_2(x)} - \mathrm{I_1(x)}|$$

$$\mathrm{B(x)} = \begin{cases} 1 & \text{if } \mathrm{D(x)} > \tau \\ 0 & \text{otherwise} \end{cases}$$

3.3 Structural similarity measure

The structural similarity measure index (SSIM) was developed by Wang et al. [8] as an attempt to measure the errors between original and compressed images in a manner that is more aligned to perceived quality than merely using MSE or PSNR. This paper describes a novel attempt to use the structural similarity measure as a difference measure that is insensitive to illumination change. The SSIM technique uses three components: luminance, contrast and structure. If we define $X = x_i | i = 1,2,...,N$ and $Y = y_i | i = 1,2,...,N$ to be the template and the test images,;[., then the SSIM can be calculated from the three components as follows

$$SSIM(x,y) = [l(x,y)]^{\alpha}.[c(x,y)]^{\beta}.[s(x,y)]^{\gamma}$$

$$s(x,y) = \frac{\sigma xy + C}{\sigma_x \sigma_y + C}$$

where α, β and γ are parameters used to adjust the relative importance of the three components, σ is the standard deviation and C is a very small constant included to avoid instability (For $l(x,y)$ and c(x,y) see [8]). In this implementation the structure component is given a high weighting (γ = 1, $\alpha = \beta = 0.1$) as structure is predicted to be least affected by illumination change and so the most useful.

The test image and template are inputted into the SSIM measurement system and the similarity value map (called by the authors SSIMmap) is outputted. The SSIMmap can then be thresholded to obtain a binary change mask. Results using the SSIM technique can be found in Section 4.

3.4 Shading model approach

In this section, we propose an image change detection algorithm (cf Figure 2) based on the shading model method proposed by Skifstad and Jain [9]. Shading models assume the intensity at a given point I_p is the product of the illumination I_i and a shading coefficient S_p.

$$I_P = I_i S_p$$

An image change detection approach based purely on shading coefficients would be highly accurate; however this is not possible because the shading coefficient cannot be determined from the image alone. Instead, the change rather than absolute value of shading coefficient between two frames is found.

Initially candidate pixels are found. Candidate pixels are defined as those that have a large absolute difference between the template and test image. The removal of the non-candidate pixels allows the algorithm to run more quickly. For each of the candidate pixels the ratio of intensities of the pixel region are calculated. The actual ratio of the individual pixel is also calculated. The difference between the pixel ratio and the block ratio is considered the likelihood of the pixel being part of an object.

$$E\left\{\sigma_s^2\right\} = E\left\{\frac{1}{N}\sum_{x\in A_i}\left(\frac{I_{x1}}{I_{x2}} - \mu_i\right)^2\right\}$$

where $E\left\{\sigma_s^2\right\}$ is the expected value of the variance and μ is the average value of the ratio of intensities. A is the region around the pixel of interest. This region must be large enough to contain sufficient intensity information around the pixel (a 10 x 10 region was experimentally found to be optimal). This variance is then thresholded to form the binary change mask.

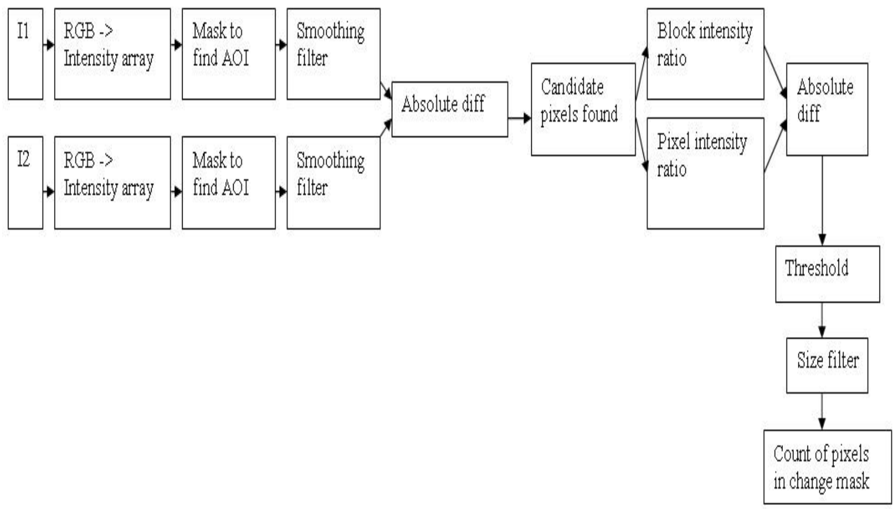


Figure 2: Block diagram of information flow in shading model algorithm

3.5 Shading model with colour data

The shading model is designed to perform well at the edges of shapes but it does not take into account colour information. For this reason an improvement to the method is suggested. The colour channels (RGB) are separated. The difference between the channels in the two frames, rd, gd, bd, are compared and the maximum difference, d, found. This maximum is normalised by subtracting the average of the other two differences.

$$rd(x) = |r(x)_1 - r(x)_2|$$
$$gd(x) = |g(x)_1 - g(x)_2|$$
$$bd(x) = |b(x)_1 - b(x)_2|$$
$$d(x) = \max(rd(x), gd(x), bd(x))$$
$$\mathrm{dnorm(x)} = \mathrm{d(x)} - \left(\frac{\mathrm{rd(x)+gd(x)+bd(x)-d(x)}}{2}\right)$$

This method uses the assumption that any illumination variance affects each of the colour channels by a similar amount.

3.6 Post-processing

The previously described methods provide the binary change mask. From this mask the system must decide on the presence of an object in the interior. The first post-processing step is a size filter. The 1's in the binary mask are clustered together (using 8 connectivity). The clusters that have less than a certain number of pixels (an experimentally found threshold) can then be removed. The results have been generated using a value of 20. This removes many of the isolated pixels that have been detected erroneously. The total number of remaining pixels is then used as a measure of how much the scene has changed and how likely it is that there is an object present.

4. RESULTS

The results shown in Table 2 are based on 80 test images (20 empty with same lighting, 20 with different lighting, 20 with people and 20 with bags). The images are in JPEG format (480 by 640 pixels).

For a complete view of the performance of the algorithms, the results will be portrayed numerically and visually. Numerical results are computed based on the following objective measures:
True positives (TP) - changed pixels correctly detected.
False positives (FP) - unchanged pixels incorrectly flagged as changed.
True negatives (TN) - unchanged pixels correctly detected.
False negatives (FN) - changed pixels incorrectly flagged as unchanged.

The percentage of correctly classified pixels (PCC) and the Jaccard Coefficient (JC) [10] are used as measures of performance.

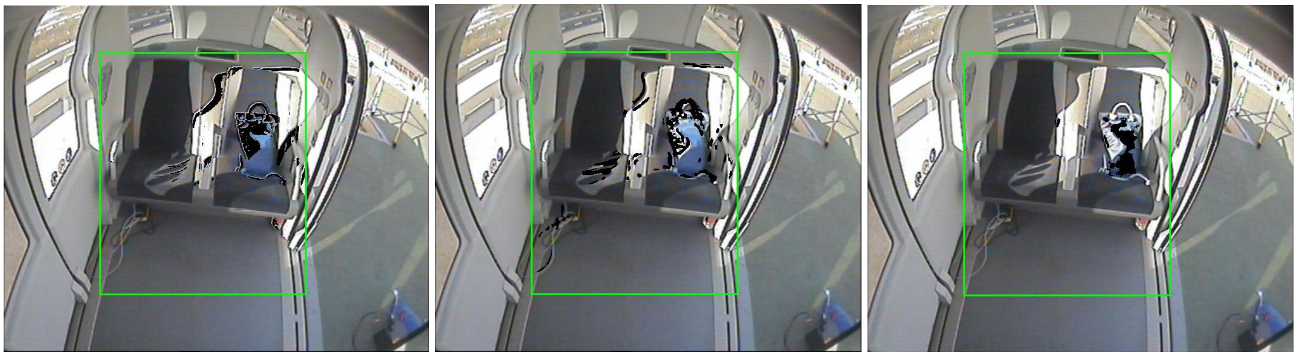
$$PCC = \frac{TP+TN}{TP+FP+TN+FN} JC = \frac{TP}{TP+FP+FN}$$

The PCC results are shown in Table 1. They show the improved shading model to be the most accurate. All three approaches are better than simple differencing particularly when the illumination in the scene has changed significantly. The SSIM method scores well but only in the case of the person identification does it have a higher accuracy than the shading model. The Jaccard coefficient (JC) is designed to put emphasis on correctly identifying objects. The results, summarised in Table 2, support the PCC with the improved shading model scoring highest overall. The SSIM suffers from identifying many shadows as objects.

The numerical results are supported by the visual inspection of the results which shows that the shading model handles shadows the most efficiently (cf. Figure 3). The results show a mean computational time of 0.4 sec.

5. CONCLUSIONS

The fast detection of people and objects has been demonstrated using a method that only requires two frames of data and one reference template. To the best of our knowledge SSIM has not been used in object detection and it is interesting to see that it performs better than simple differencing.



(a) (b) (c)

Figure 3: Results of various image change detection methods. (a) Simple differencing method using difference threshold of 0.8. (b) SSIM method using threshold of 10. (c) Shading method with colour information using difference threshold of 0.8.

Scenario	Absolute Differencing	SSIM	Shading Model	Shading Model Adaption
Empty Scene (same lighting)	100.0	100.0	100.0	100.0
Empty Scene (different lighting)	88.6	96.0	99.5	99.4
Person	70.8	70.9	68.4	70.2
Bag	94.7	94.3	95.2	96.9

Table 1: Image change detection success rate results using PCC method

Scenario	Absolute Differencing	SSIM	Shading Model	Shading Model Adaption
Empty Scene (same lighting)	100.0	100.0	100.0	100.0
Empty Scene (different lighting)	88.6	96.0	99.5	99.5
Person	68.0	69.0	68.1	69.4
Bag	94.6	94.2	95.2	96.8

Table 2: Image change detection success rate results using Jaccard Coefficient method

The shading model demonstrates the best performance of the methods studied and is further improved by the addition of the colour differencing technique. Further experimentation of the algorithm in a working PRT system will enable more detailed performance conclusions to be drawn. It is hoped this technique will allow the PRT system to run efficiently with reduced human supervision.

REFERENCES

[1] M Lowson, “New approach to effective and sustainable urban transport,” *Transportation Research Record*, , no. 1838, pp. 42–49, 2003.

[2] P.J Escamilla-Ambrosio, “A multiple-sensor multiple-target tracking approach for the autotaxi system,” *IEEE Intelligent Vehicles Symposium*, pp. 601 – 606, 2004.

[3] P. J. Escamilla-Ambrosio and N Lieven, “Fuzzy logic obstacle identity declaration and fusion in the autotaxi system,” *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1–6, 2007.

[4] R. J. Radke, S. Andra, O. Al-Kofahi, and B Roysam, “Image change detection algorithms: a systematic survey,” *Image Processing, IEEE Transactions on*, vol. 14, no. 3, pp. 294–307, 2005.

[5] X. Dai and S Khorram, “Requirements and techniques for an automated change detection system,” *Geoscience and Remote Sensing Symposium Proceedings, IGARSS ’98. IEEE International*, vol. 5, pp. 2752–2754, 1998.

[6] P. R. Coppin and M. E Bauer, “Digital change detection in forest ecosystems with remote sensing imagery,” *Remote Sensing Reviews*, vol. 13, no. 3, pp. 207–234, 1996.

[7] A Singh, “Review article digital change detection techniques using remotely-sensed data,” *International Journal of Remote Sensing*, vol. 10, no. 6, pp. 989–1003, 1989.

[8] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.

[9] K. Skifstad and R Jain, “Illumination independent change detection for real world image sequences,” *Comput. Vision Graph. Image Process*, vol. 46, no. 3, pp. 387–399, 1989.

[10] P. L. Rosin and E Ioannidis, “Evaluation of global image thresholding for change detection,” *Pattern Recognition Letters*, vol. 24, pp. 2345–2356, 2003.