Technical Report

Project Kitchen Occupation

TSBB11 HT 2013

Version 1.0

Status

Reviewed	–	2013–12–13
Approved		

2013–12–13

Project Kitchen Occupation

Bilder och Grafik CDIO, HT 2013
Department of Electrical Engineering (ISY), Linköping University

Participants

Name	Tag	Responsibilities	Phone	E-mail
Mattias Tiger	MT	Project manager	073–695 71 53	matti166@student.liu.se
Erik Fall	EF	–	076–186 98 84	erifa226@student.liu.se
Gustav Häger	GH	System integration	070–649 03 97	gusha124@student.liu.se
Malin Rudin	MR	–	073–800 35 77	malru103@student.liu.se
Alexander Sjöholm	AS	–	076–225 11 74	alesj050@student.liu.se
Martin Svensson	MS	Documentation	070–289 01 49	marsv106@student.liu.se
Nikolaus West	NW	Testing	073–698 92 60	nikwe491@student.liu.se

Homepage: TBA

Customer: Joakim Nejdeby, Linköping University, Origo 3154
Customer contact: 013–28 17 57, joakim.nejdeby@liu.se
Project supervisor: Fahad Khan, Linköping University, fahad.khan@liu.se
Examiner: Michael Felsberg, michael.felsberg@liu.se

Contents

List of Figures

Document history

Version	Date	Changes	Sign	Reviewed
0.1	2013–12–13	Initial draft	MS	MT

1 Introduction

There exists many places in society where the degree of human occupancy and movement flow is desirable to know as basis for decision making. Such data answers if it is necessary to build more rooms and provides knowledge of actual user or consumer patterns. Example usages are measuring resource usage of public spaces, or which part of a store that attracts most people. It provides vast opportunities in resource management, marketing, sales and scheduling. There exist some plausible solutions to estimating the number of people at a location such as using cell phones or motion detectors, but this project aims at an image based approach with the possible benefits of being both cheaper and more robust.

1.1 Background

Today Linköping University has many places with similar functionality, e.g. student kitchens where students are provided with the ability to warm food brought with them. Linköping University has several such kitchens all over its campuses. Critics claim that there are too few student kitchens with microwave ovens and that the existing ones usually are overcrowded. That all kitchens are overcrowded at the same time has not been confirmed by sample inspections. One standing hypothesis is that students don't know where all the kitchens are nor that they want to risk going to a kitchen in another building in case that is full as well.

Linköping University has an ongoing project with the purpose of enabling the students to see the usage of some of the schools resources (e.g. group rooms) online. The aim of this project is to supply that system with data regarding the usage of student kitchens. It will provide all students with the ability of visualizing the crowdedness of each kitchen, thus providing them with the means of finding the closest, least occupied kitchen available.

1.2 About this document

This docment, together with the code reference manual (appendix C) constitutes the final documentation of the project and describes the current implementation together with some failed attempts and conclusions drawn from these. Performance results of the current implementation are presented in section 4 of this document.

2 System description

The text about entire system stuff and more stuff

Figure 2.1: *System overview.*

2.1 Config file system

The system configuration file contains all image processing pipeline parameters together with the address of the current destination web service, where counting results are to be presented.

2.2 Network module

The network module manages all system input/output. Currently, support exists for all OpenCV-compatible RGB cameras as well as the Microsoft Kinect depth sensor. The module is designed to be as modular as possible, allowing for a relatively easy replacement integration of new sensor types. Support also exists for running several sensors in parallel. Communication with the web server also takes place here.

2.3 Algorithm Interface

Algorithm interfaces and stuff

2.4 Debug and Config GUI

short about config gui

Figure 2.2: *System overview.*

3 Current pipeline

What is this?

3.1 Image processing

imgproc stuff

3.2 Tracking

tracking stuff

3.3 Queue detection

queue stuff

4 System Evaluation

One of the most important parts of a project is the ability to evaluate it. Without it there is no way to compare results with others... more to come

4.1 Ground Truth

TODO: info on acquisition of ground truth: The evaluation needs access to some sort of ground truth which is defined as the best possible achievable tracking output. Throughout this part of the document an object is defined as a position by the ground truth and a hypothesis as the output from the tracker. The method only allows one-to-one correspondence between objects and hypothesis and in case of conflict the combination yielding the lowest total distance error is chosen. The ground truth data and video clips used for evaluation are all from the CAVIAR project [7].

4.2 MOTA

Info about tracker evaluation MOTA is a measure of accuracy with respect to how many mistakes are made by the tracker. It consists of four variables: misses, false positive, mismatches and number of objects. A miss is when no hypothesis suggested is close enough to an object, close enough being defined by a threshold, T. This is the only design parameter in this method. If the distance between an object and the closest hypothesis is larger than T, the object is considered a miss. If a hypothesis has no object within threshold distance it results it is labeled a false positive. One important feature of a tracker is the ability to keep objects identities correct. If this is not the case and an object is changing identity between frames one mismatch error is added for every change. The number of objects variable is defined as the total number of objects trackable according to the ground truth in the current frame. For a given image sequence, equation (**??**) can be used to compute the MOTA.

4.3 People count

Info about current evaluation metrics.

5 Results

In this section the results of the project is summarized for some sequences chosen together with the other group. The sequences, as well as the ground truth used are provided by the EC Funded CAVIAR project, [7].

5.1 Evaluation Scores

In table 5.1 below the video clips chosen for the evaluation are presented, along with the achieved accuracy measurements. TODO: Fix table:

Sequence Name	*In*	*Out*	*Count*
training data	0.56	5.01	5.01
evaluation data	0.70	5.30	5.01
other data	0.65	5.09	5.01
data stuff	0.90	5.44	5.01

Table 5.1: *Counting performance according to the evaluation metric as described in section 4.*

5.2 Discussion

Discussion stuff and more stuff.

6 Conclusions

6.1 Hardware

Conclusion about stuff

6.1.1 Possible improvements

possible improvements to said stuff

6.2 Software (or something else)

Conclusion about stuff

6.2.1 Possible improvements

possible improvements to said stuff

6.3 Image processing

Conclusion about stuff

6.3.1 Possible improvements

possible improvements to said stuff

6.4 Tracking

Conclusion about stuff

6.4.1 Possible improvements

possible improvements to said stuff

6.5 Queue

Conclusion about stuff

6.5.1 Possible improvements

possible improvements to said stuff

A An unsuccessful attempt - people detection using background subtraction.

Some group members had earlier experiences of tracking using background subtraction and therefore the first approach of the development was trying that method. Their approach was to generate a binary image from the raw image using [1] and after that use OpenCV's built in functions boundingRect in combination with own developed functions to separate e.g. blobs of humans. This approach was developed and performed good for tracking people moving relatively fast, see figure A.1, but worse or not at all for people moving slow and standing still. The tracker should be able to track a queue, which moves slowly or stands still, and in this regard the method using background subtraction ended up being too bad. The problem was that still or slow objects melted into the background and therefore disappeared from the binary image, different learning rates were tested but without good results. This is shown in figure A.2. The nature of the binary image makes it hard to differentiate people standing close to each other, because it then creates one united blob. The algorithm described in [1] was chosen since it was the most recent developed background subtractor in OpenCV, and also the best performing one according to the OpenCV reference manual.

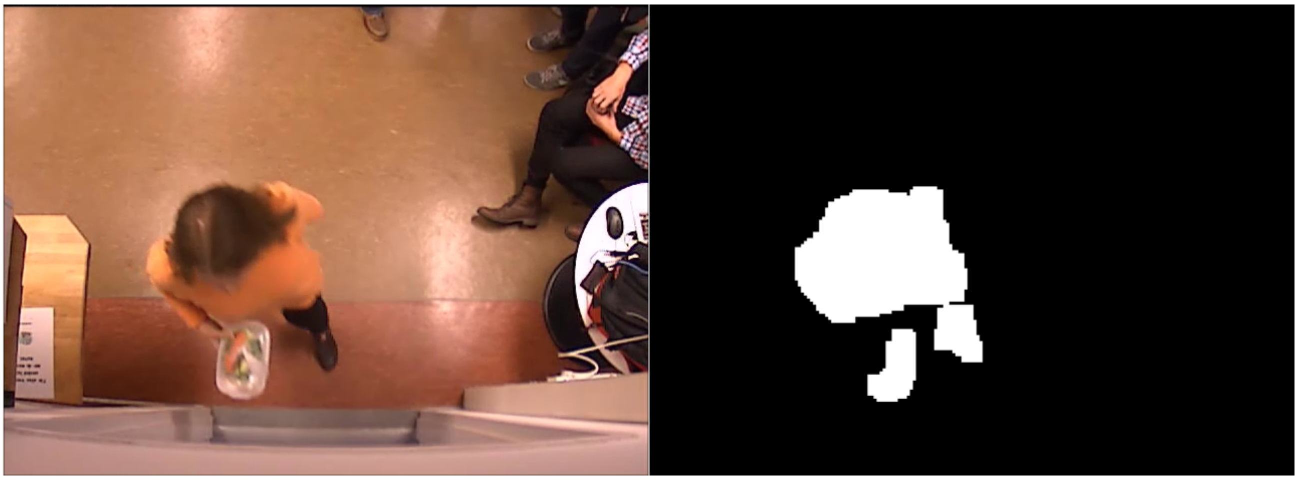

Figure A.1: *The left image shows a moving person who is detected by the background subtractor. The right image shows the binary image generated from the left image.*

Figure A.2: *The left image shows a persons who is not detected by the background subtractor, because of insufficient movement. The right image shows the binary image generated from the left image.*

B An unsuccessful attempt - head detection by means of circle detection.

In [1] Gardel et.al. a method of head detection from above based on circle detection is proposed. In their approach circles, i.e. heads, are detected by first performing canny edge detection on each image frame and then performing a form of hough voting to detect circles. The hough voting is done by combining the results of a series of different filters designed to give high responses at circle and ellipse centers. We implemented their approach and experimented with many different variations of filter sizes and canny variables, with both types of circle filters described in the paper. We also tried using background subtraction on the canny image, using a mixture of Gaussian-based foreground segmentation [2] of the raw image, which gave slightly better results in simple situations. We were however not able to get any usable results for anything but the simplest cases, and therefore gave up on this approach after many fruitless attempts at tuning the system. Outputs from some of the different steps of cases where the approach was both successful and unsuccessful can be seen in figure B.1 and B.2 below.

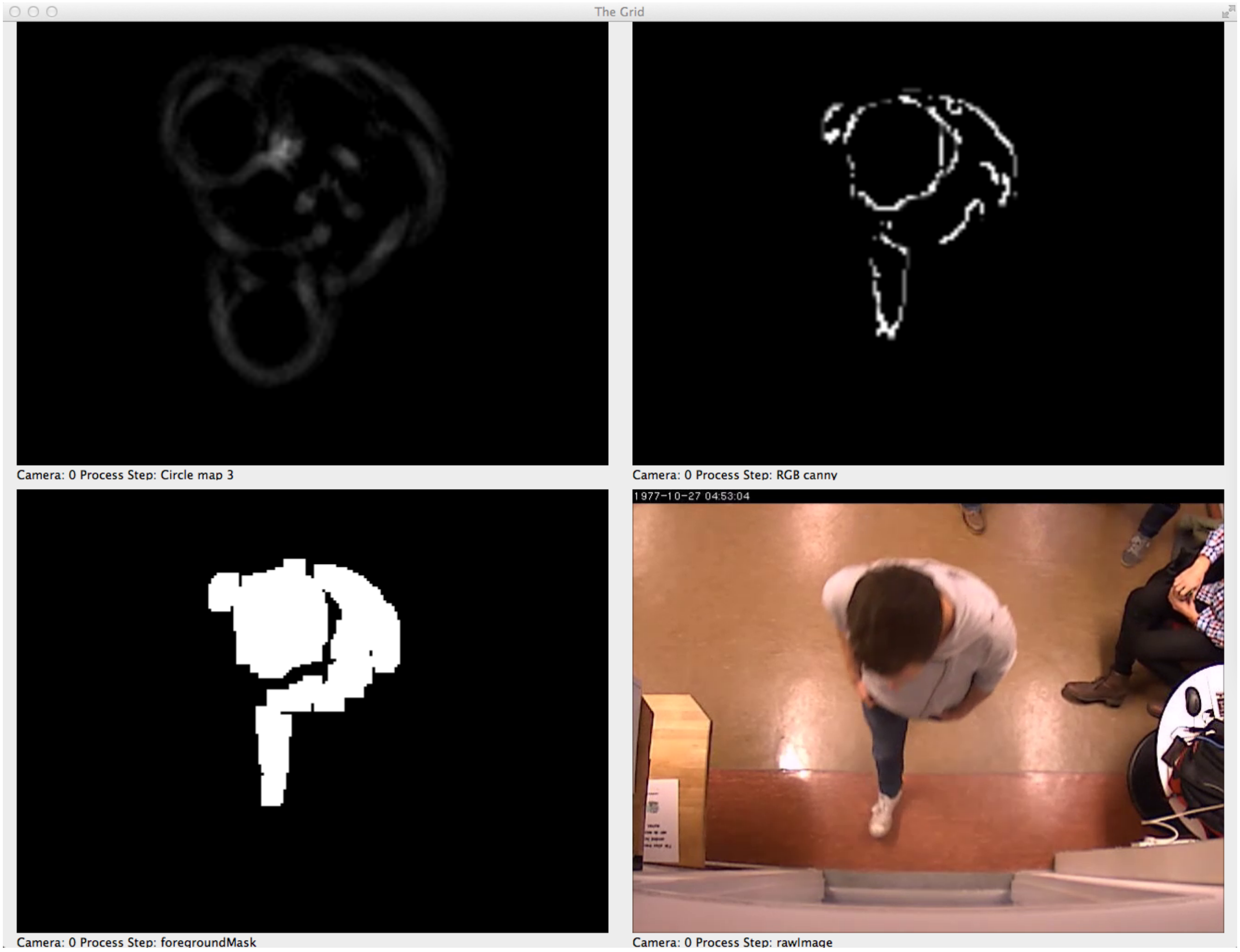

Figure B.1:
Top left: output from one of the circle filters.
Top right: output from the canny edge detection after background subtraction.
Bottom left: foreground mask.
Bottom right: raw image.

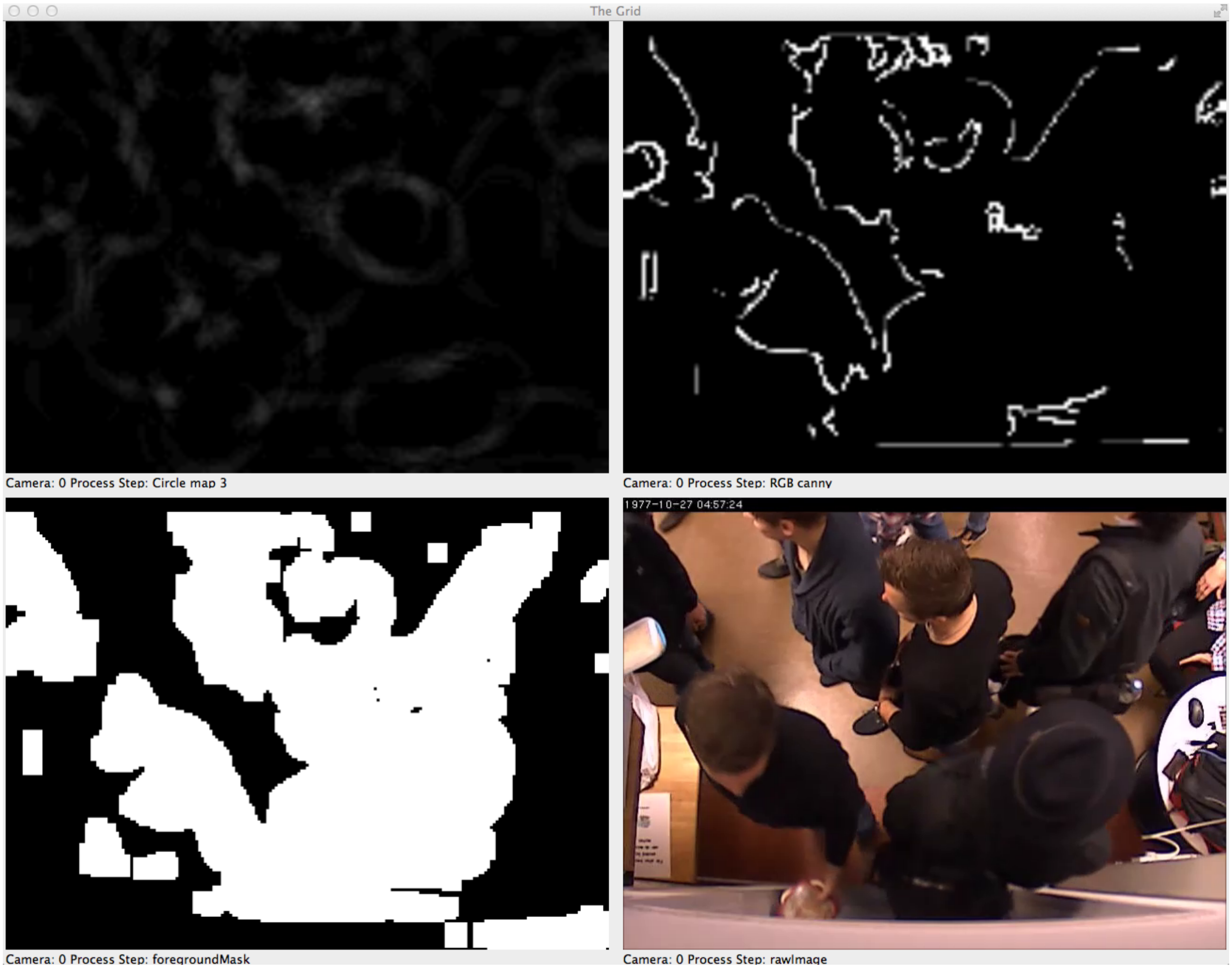

Figure B.2:
Top left: output from one of the circle filters.
Top right: output from the canny edge detection after background subtraction.
Bottom left: foreground mask.
Bottom right: raw image.

References

[1] Gardel, A., Bravo, I., Jimenez, P., Lazaro, J.L. & Torquemada, A.
"*Statistical Background Models with Shadow Detection for Video Based Tracking*,"
Intelligent Signal Processing, 2007. WISP 2007. IEEE International Symposium on?? Page: 1-6.

[2] Zivkovic, Z. & Heijden, F.
"*Efficient Adaptive Density Estimation per Image Pixel for the Task of Background Subtraction*,"
Pattern recognition letters, Vol. 27, No. 7. (2006), pp. 773-780.

EXAMPLE REFERENCES ONLY, REMOVE BEFORE HANDING IN

[3] Sonka, M., Hlavac, V. & Boyle, R. *Image Processing, Analysis, and Machine Vision.*
Toronto: Thompson Learning, cop. 2008, 3rd ed., ISBN 0495244384.

[4] Wood, J. (2007) "*Statistical Background Models with Shadow Detection for Video Based Tracking*,"
Master thesis, Linköping University, Department of Electrical Engineering.

[5] Gustafsson, F., Ljung, L. & Millnert, M. *Signal Processing.*
Studentlitteratur, Lund, Sweden, 2011, 1st ed., ISBN 978–91–44–05835–1.

[6] Bernardin, K. & Stiefelhagen, R (2008)
"*Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics*,"
Interactive Systems Lab, Institut für Theoretische Informatik,
Universität Karlsruhe, 76131 Karlsruhe, Germany

[7] "*CAVIAR: Context Aware Vision using Image-based Active Recognition*,"
EC Funded CAVIAR project/IST 2001 37540
http://homepages.inf.ed.ac.uk/rbf/CAVIAR/