

Problem Solving and Engineering Design part 3

ESAT1A1

*Max Beerten
Brent De Bleser
Wouter Devos
Ben Fidlers
Simon Gulix
Tom Kerkhofs*

Counting and recognizing non-moving objects by means of image processing

PRELIMINARY REPORT

Co-titular
Tinne Tuytelaars

Coaches
Xuanli Chen
José Oramas

ACADEMIC YEAR 2018-2019

Declaration of originality

We hereby declare that this submitted draft is entirely our own, subject to feedback and support given us by the didactic team, and subject to lawful cooperation which was agreed with the same didactic team.

Regarding this draft, we also declare that:

- 1. Note has been taken of the text on academic integrity
(<https://eng.kuleuven.be/studeren/masterproef-en-papers/documenten/20161221-academischeintegriteit-okt2016.pdf>).*
- 2. No plagiarism has been committed as described on <https://eng.kuleuven.be/studeren/masterproef-en-papers/plagiaat>.*
- 3. All experiments, tests, measurements, ..., have been performed as described in this draft, and no data or measurement results have been manipulated.*
- 4. All sources employed in this draft – including internet sources – have been correctly referenced.*

Contents

| | |
|---------------------------------------|-----------|
| Contents | 3 |
| List of Figures | 4 |
| 1 Introduction | 1 |
| 2 Problem Description | 1 |
| 3 Design | 2 |
| 3.1 Hardware | 2 |
| 3.2 Software | 3 |
| 3.2.1 Analysis RGB image | 3 |
| 3.2.2 Analysis Depth Sensor | 6 |
| 4 Implementation | 8 |
| 5 Further planning | 10 |
| 6 Budget management | 10 |
| 7 Course Integration | 10 |
| 8 Conclusion | 11 |
| 9 References | 12 |
| 10 Appendix | 13 |

List of Figures

| | | |
|---|---------------------------------------------------------------------------------------------------|----|
| 1 | The example included in the assignment. | 2 |
| 2 | A comparison between the 3 different methods. | 5 |
| 3 | A comparison with the use of filters. | 6 |
| 4 | The original RGB image (left) and the image after the Sobel-Feldman operator (right) | 7 |
| 5 | The original image after using a Sobel-Feldman operator and a threshold filter | 8 |
| 6 | Problem with stopping criteria Moore-Neighbor tracing algorithm. | 10 |

1 Introduction

Digital image processing has been a crucial part of the current digitalisation movement. From industrial machinery to customer amusement, the vision of computer-aided systems has become a given for most users. While image alteration and manipulation remain a core part of this image processing, nowadays other image related problems are being solved by artificial intelligence. Most were considered to be an important part of digital image processing. Among these, the problem of this paper can be found: feature extraction. The ability to count objects in an image to be more exact. So why use 'traditional' methods to solve this problem? While being a great way for unravelling many problems, artificial intelligence mostly provides general solutions. However, certain cases are solved more efficiently by specific schemes. Such is the case with object counting: while deep learning algorithms need a big data set as training material, standard image processing only requires the image itself.

Regardless which way a method processes images, it needs a visual source. In this paper the focus is on live object counting, which is only possible with a camera. Evidently, the choice of hardware greatly impacts the methods that can be used. This choice will be covered in !TITEL HARDWARE HERE!.

By far the most important part of this task is the algorithm by which the items in the picture will be numbered. Classically, object counting algorithms have a standard group of steps: filtering, converting to an intensity matrix, edge detection, converting to a binary matrix, boundary boxing and the counting itself. These segments don't have a fixed order and can occur multiple times in the final method. Most of these steps can also be approached in different ways. A wide range of possible filters, kernels, edge detection methods, etc. exist, which all have their benefits and drawbacks.!REFERENTIE BOEK! These choices will be discussed in !TITEL SOFTWARE HERE!.

These methods, while being the core of the solution, are fairly simple to implement with the use of libraries or built-in functions. In this paper is opted to give a full implementation of these functions, limiting the usage of libraries to the minimum, in !TITEL IMPLEMENTATION HERE!. If the functions are deemed to be basic, only a simple explanation will be given.

2 Problem Description

The object counting system described in this report is capable of counting non-moving objects in a basket. These objects can vary in shape, size and colour. Thus, the colour of both the basket and its contents are free from restrictions.

In the primary stage of this paper, not all these variables are taken into account. The simplest objects, which the system is required to count, are rectangles, cylinders and circles, all with a uniform colour. If possible, the circumference of these objects can be outlined and measured as shown in Fig. 1.

All of this is done in real-time and with a budget of €250.



Figure 1: The example included in the assignment.

3 Design

3.1 Hardware

The hardware to create a system as described above, is not complicated. In essence, it consists of a computer, a camera and a cable to transfer the data between the prior named necessities. Each of these hardware components is discussed in the following section.

Choosing the camera is a vital element in this project. If chosen poorly, it can fiercely limit the outcome of the final algorithm. There are three main options for visual input: an ordinary webcam, an industrial camera or a camera with built-in depth sensors. Each with its pros and cons. A webcam is cheap and readily available but does not assure good image quality and easy access to its data. A camera for industrial usage is rather expensive, especially with a budget of €250. Industrial grade options which are cheap enough exist, but these models deliver their images in greyscale. This greatly limits the possible methods which can be used. Thirdly, the depth sensing cameras are available in a reasonable price range and deliver, overall, good quality data. Moreover these models have the added benefit of depth sensor which, in contrast to the previous option, adds more possible ways to solve the problem.

Having considered all of the above, the best option is the latter one. More specifically, the system described in this report is based on a Kinect V2 from Microsoft. This camera has a color lens with a resolution of 1920 by 1080 pixels and a corresponding field of view of 84.1° by 53.8° (Smeenk, 11 Mar 2014). The high resolution ensures an accurate matrix representation of the real image. Each color frame pulled from the Kinect V2 is represented by an array structure of 1080x1920x3. Every element corresponds with a pixel of the image and varies between 0 and 255. Obviously it can be separated into three different matrices each belonging to \mathbb{R}^2 and based on a different colour: red, green or blue.

Next to the colour camera, the Kinect also possesses a depth sensor. An infrared projector

and camera make this possible(Jiao, Yuan, Tang, & Wu, Nov 2017). It provides a 424x512 array making the depth image one of roughly 200000 pixels. The field of view of this function is 70.6° by 60°. Note that the depth camera provides data about parts of the environment that the color camera does not see, and vice versa. When the computer reads the depth data, every number in the matrix represents a distance in millimetres. Obviously there are some restrictions. This technology only provides correct information if the object is at a distance located in between half a meter and 4 meters. This has to be taken into account for further implementation of this paper.

As second element of hardware the computer has a less important role. Preferably, OSX isn't used as operating system for this application because the Kinect drivers do not exist for Macintosh computers. If the reader has a Mac, problems can be avoided by running either Windows or Ubuntu via a virtual machine. The algorithm should run in an acceptable time frame on every machine.

To conclude this section a brief word on the necessary transfer cable. Since a depth sensing camera is used, two types of data (depth and color) need to be transferred. The Microsoft OEM Kinect Adapter makes this possible. The special adapter is the only available option and consists of two general parts. One part for delivering current to the camera and the other to transfer both types of data to the connected computer.

3.2 Software

There are a lot of options when it comes to software and a wide range of different algorithms for image processing exist. The diagram on attachment C shows a couple of different methods. There is no 'right way' to count objects in an image.

3.2.1 Analysis RGB image

Analysing RGB images is a creative process. Different approaches have their own advantages and disadvantages. The only general ideas that are common throughout most algorithms are:

- Converting the RGB image to greyscale
- Run filters over the image to remove noise

In the next scope, three general methods are featured and briefly discussed. Each was investigated in prospect of this paper.

Method 1

This method is the most simple and straightforward to implement. As input it requires a filtered greyscale image. This is passed through a thresholding algorithm with a pre-defined threshold value. The output is a binary matrix. Thus, this array only has 0's and 1's as elements, respectively representing the colours black and white. The key to solving the problem in this specific scheme is writing code that finds the threshold value based on environmental parameters. When in possession of a truly black and white image, a simple edge detection program is run which makes the edges visible.

Advantages: It's an easy and fast algorithm.

Disadvantages: With a pre-defined threshold value it just classifies pixels based on colour. A dynamic value is required.

Method 2

The second method tackles the colour analysis in the opposite order than the first method, as it commences with an edge detection algorithm. Since the input image is still very complex, it is first converted to a filtered greyscale image. Still, this edge detection is more comprehensive than that one from method one. The output still is a greyscale image, contrary to the binary array the reader might expect. This is followed by some thresholding code with a pre-defined threshold value. The current image is now represented by a matrix where the edges are outlined using binary elements.

Advantages: It detects all kinds of objects, not based on colour or shape.

Disadvantages: The boundary between different objects needs to be clear for this to work.

Method 3

The third way takes a different approach to solving the analysis of the colour image. When using this, a compromise in functionality is made. Since it needs a picture of the empty background without any objects, the user experience is worsened. After getting a background image, the picture of the situation with objects gets filtered and the algorithm converts it into a greyscale image. Using this less complex matrix, the code loops through the image pixel by pixel. This necessary but time consuming loop checks if the pixel on the image is more or less the same as the corresponding pixel on the background image. If located within a pre-determined range, that element of the array gets classified as background. The consequence is that the output is a binary image with clear-cut objects.

Advantages: It is very good in detecting objects, not being based on colour or shape.

Disadvantages: There needs to be an image of the empty background. Note that the lighting conditions have to be unaffected in between taking the needed pictures for this algorithm.

Choice of method

After comparing these different schemes, the second method comes out as the better of the three. See Fig.2 for the comparison.

The first step, as discussed above, is to convert the image to greyscale (*Greyscale*, n.d.). This is easily done by calculating a weighted average of the values of all three red, green and blue matrices as shown in the following equation. A linear combination is made.

$$greyscale_image(row, col) = 0.2989 * RED + 0.5870 * GREEN + 0.1140 * BLUE \quad (1)$$

The used coefficients count up to 1, so the values in the greyscale image can vary from 0 to 255. All these values $greyscale_image(row, col)$ form the new image.

Before running the image through an edge detection algorithm, two filters are applied. Both blur the image to an extent such that noise after edge detection is considerably reduced. This effect is visualised in Fig. 3 Firstly a Gaussian blur (*Kernel (image processing)*, n.d.) is applied. Most filters are a convolution of a kernel with the image. For a Gaussian blur the G kernel found below is used. This is just a weighted average. This has as consequence that the elements centered around the main pixel have bigger weights than those at the edges.

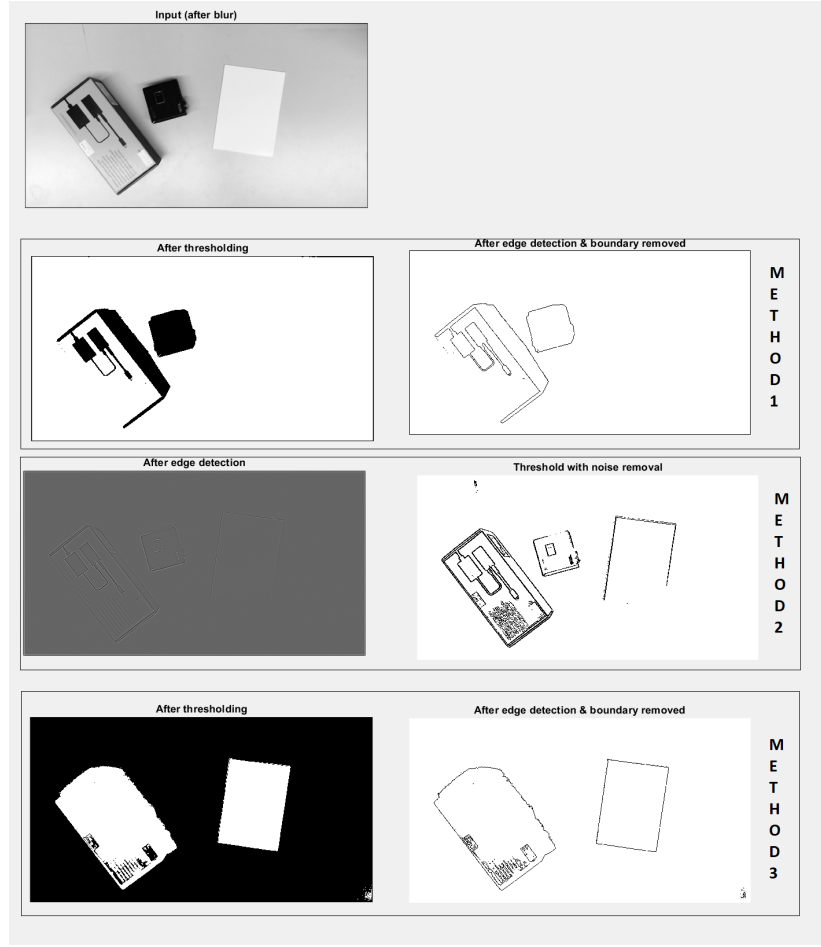


Figure 2: A comparison between the 3 different methods.

$$G = (1/159) * \begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix} \quad (2)$$

The second blur is a mean blur (R. Fisher & Wolfart, n.d.-b). This is almost the same, the actions are just done with a different kernel. This kernel calculates the average of the values around the pixel.

$$M = (1/9) * \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (3)$$

Note that both G and M have a norm of 1. If this wasn't the case, pixel values of the filtered image could exceed the boundary values of 0 to 255 for uint8 numbers.

After both filters are applied, the image is ready to be run through an edge detection algorithm(R. Fisher & Wolfart, n.d.-a). This algorithm is on itself also a filter with ker-

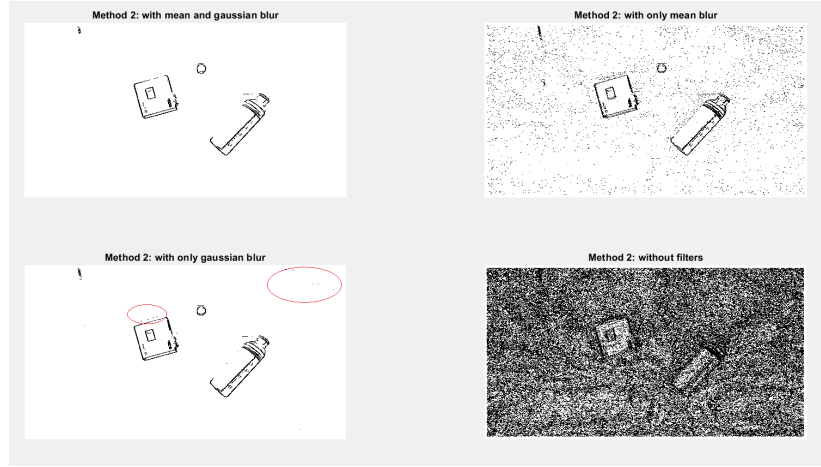


Figure 3: A comparison with the use of filters.

nel done by the matrix L given below. It calculates the *spatial derivative* or in simpler words, it highlights regions of rapid intensity change.

$$L = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad (4)$$

Note now how the kernel uses the pixels next to the evaluated pixel to see how much intensity changes. If the image wouldn't have been filtered before convolution with L , more 'edges' would have been drawn because more unregular areas exist.

Note also how this convolution returns a new image which can have negative values for its pixels. The more negative the value, the darker the image. A different number type is used. The threshold algorithm, used in the following step, is based on this feature. This algorithm runs through to whole matrix and assigns each value with either a 0 or a 1. It decides this by assessing if the current value is either smaller than or bigger than a threshold value, respectively. After conducting multiple experiments and testing, a threshold value of 2 seems to do the trick. A more dynamic way of determining this value may be developed in the next weeks. After applying this edge detection, the matrix becomes a binary image with only the edges in white. Based on these edges it is possible to outline the objects and count them, but further research and programming has to be done to complete the whole program.

3.2.2 Analysis Depth Sensor

Using only the RGB image does have some shortcomings. It is rather difficult to distinguish an object from its shadow, a multicoloured object could be seen as multiple different objects and a lot of reflection could make an object undetectable. These are some of the reasons why enrichening the object counting algorithm with the usage of a depth sensor is advised. Like featured in the section about the hardware, each element of the input data represents a distance in millimeters.

Firstly the code should be able to provide a clear difference in height between the objects and the background using the depth data. This is followed with a filter to get rid of the existing

noise reduction. At last, the filtered matrix will be used to detect the edges of the objects and thus detect the items themselves. **The code that accompanies this description, can be found at page...**

Detection of the difference in height

The goal is to see a clear difference between the objects and the background. This can be achieved in different ways: it is possible to use a threshold and label everything closer than this predetermined distance as an object. A disadvantage of this method is that this value will be different for different vertical positions of the kinect v2. Also, the image of the sensor contains some noise. For example: a picture of a big flat table will not be viewed as a equidistant surface. The elements of the matrix will be different. Another, and more preferred, method would be to use a Sobel-Feldman operator (Sobel, 2014). This operation approximates the gradient in each of the points of the matrix, and gives an idea where there is a sudden difference in height (thus where there might be an object). It works by convolving 2 kernels with the image matrix A to become G_x and G_y : respectively one for the horizontal and one for the vertical change in height:

$$G_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * A$$

$$G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A$$

$$G = \sqrt{G_x^2 + G_y^2}$$

In the last equation, G is the magnitude of the total gradient as well as the value inserted in the new matrix. The result can be seen in fig. 4.

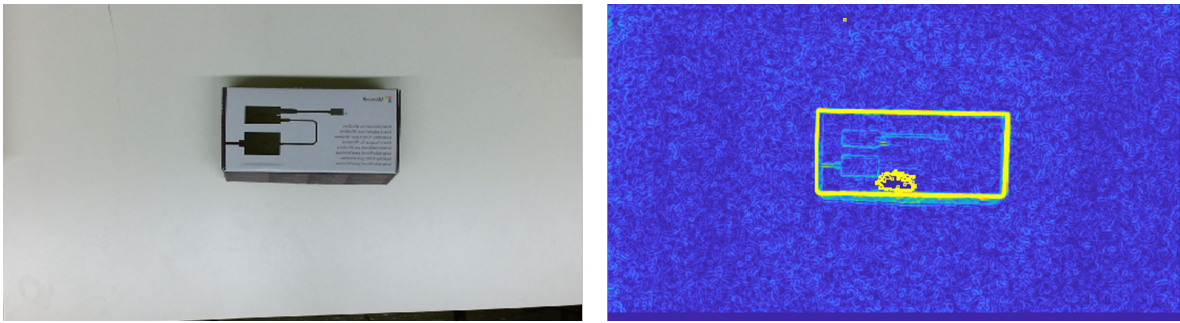


Figure 4: The original RGB image (left) and the image after the Sobel-Feldman operator (right)

Filtering of the noise

After adding all the different magnitudes of the gradients to an array, some anomalies still exist. There can be some impossible elements, like points that seem to be further away than

the basket, or fluctuations in areas that are supposed to be flat (noise). The simplest way to solve this problem would be to use a maximum and minimum threshold: The maximum threshold can be a value that is further away than the basket. These values are impossible and the corresponding values in the matrix can be set to zero. The minimum threshold can be decided by empirical research. Values lower than this value can be seen as noise and thus can be set to zero. The final result can be seen in fig. 5.

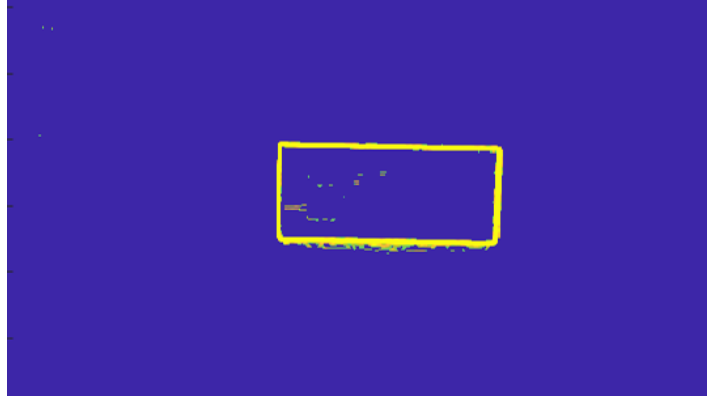


Figure 5: The original image after using a Sobel-Feldman operator and a threshold filter

4 Implementation

Throughout the project a lot of different approaches were tested and discarded. But in essence, they all do the same thing. They convert the original image to a binary image. Afterwards, this binary array is analysed and a simple algorithm suffices to count the objects. In this phase of the program the same code is applicable. This code consists of a few important parts: the actual counting and the drawing of the boundary boxes.

Counting of the objects

The central objective of this paper is counting the amount of objects in a specific rectangular field of view. The general approach to this problem is converting the image to a binary image where black pixels represent the background and white pixels represent the objects. By counting the groups of pixels, it is possible to know how many objects the original image contains. In the image processing toolbox for matlab (*Mathworks*, 2018), a few functions exist that are very usefull for this kind of tasks. One of these functions `bwlabel` actually counts group of pixels of at least 8 that are connected. The syntax of this function goes as follows:

$$[L, num] = bwlabel(BW) \quad (5)$$

where `BW` represents the binary (or black and white) image; `num` represents the number of objects in the `BW` image and where `L` represents a matrix where the first group of pixels are numbered 1, the second group 2 etc. that way it's easier to get a count for how many objects there are.

Boundary boxes

The image processing toolbox really simplifies the drawing of boundary boxes. Once a binary image is obtained the function `regionprops` can extract properties about image regions. Where image regions are defined as 8-connected components in an binary image. This means that each image region contains at least 8 interconnected white pixels, since the black pixels are registered as background. The property that's interesting for this part of the project is called 'boundingbox'. This property returns for every image region the smallest rectangle containing this region. In two dimensions this is a vector with 4 values, the x-coordinate of the upper left corner, the y-coordinate of that corner, the width and the height. The function

$$rectangle('Position', pos) \quad (6)$$

where 'Position' declares the input and where `pos` is the input obtained from `regionprops`, can easily display this boundingbox.

Edge detection

There are a lot of ways to implement edge detection. Edge detection algorithms as described in PARAGRAPH exist for greyscale image. But if a binary image is available, this becomes much easier. For starters there exists a function in the image processing toolbox called `bwboundaries`. The syntax of that function goes as follows:

$$B = bwboundaries(BW) \quad (7)$$

where `BW` represents the input, this is a binary image which only consists out of black and white pixels; and `B` represents the output, which consist out of a cell array with `N` elements (number of image regions in the binary image), all these elements contain a list of the boundary pixels. Which in turn are fairly easy to draw. They can be inserted in the matrix of the image by replacing values, this is done by looping through the cell arrays. The advantage of this method is that the image can actually be printed. When they are drawn on top of the image with a function like `visboundaries` the actual values of the pixels stay unchanged, but it become different figures. One with the image and another on top of it with the edges. The function `bwboundaries` implements the Moore-Neighbor tracing algorithm. The algorithm loops through the entire matrix until it finds a white pixel (a pixel that belongs to an image region). This pixel is defined as the start pixel. Once it finds a start pixel it searches for the next connected white pixel. This means another white pixel in one of the eight regions around the start. The algorithm does this by examining the pixels in a clockwise direction. Once it finds a new white pixel, this pixel is added to the sequence `B` and becomes our new start pixel. This process keeps on running until the algorithm visits the first start pixel for a second time. The only problem with this algorithm is that sometimes the first start pixel is visited for a second time before all of the outline is visited (See fig. 3).

This problem is resolved with the Jacob's stopping criterion. Which states that the algorithm can stop once the first start pixel is visited out of the same direction as it was initially entered. This leaves four possibilities that need to be checked, from below, from the left, from above or from the right. With this additional criteria, every pixel at the edge of a connected region is visited. To find the edges of all the interconnected image regions this process is repeated until every pixel of the image matrix has been checked.

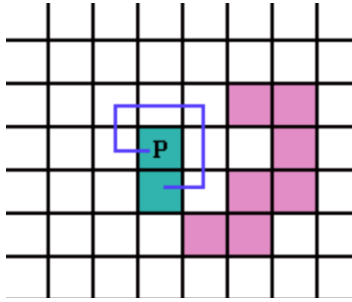


Figure 6: Problem with stopping criteria Moore-Neighbor tracing algorithm.

5 Further planning

Being halfway through the project, a visual timeframe is created. In the appendices, a Gantt chart can be found.(APPENDIX) This visualizes the current state of the solution for the described problem, as well as the schedule for the next few weeks. This project contains five milestones, two of them are already achieved. These have a minor value in prospect to the total paper though. The most important occupancy of the next weeks is implementing key elements of the final algorithm. Key components like filling the edges, creating the boundary boxes and eventually counting the number of objects still need the necessary attention. All of this while the given deadlines need to be respected. As it is possible to view in the chart, the decision to start rather early on the folder is made. A professional representation of the findings takes time. So planning it like this, ensures enough time to perfect the folder. In general, the project is on schedule. From a critical point of view, too much time writing the report during the team sessions was wasted. For the final paper, more individual work is recommended and will happen.

6 Budget management

As seen above, the system explained in this paper primarily consists of software which on its own doesn't cost anything. On the contrary, the necessary hardware is rather costly. The current set-up consists of a tripod and the electronics. The tripod is lend for free by the faculty thus the only remaining costs are the Kinect v2 and its adapter to connect with a personal computer.

With a budget of 250 EURO, this is feasible. Both the Kinect and the adapter have been ordered but as of writing this paper, a fixed price isn't known. At the current market prices, the estimated cost is €200. The remaining €50 are a safe backup for other small costs.

7 Course Integration

For this project, some courses from the first three semesters at the institute of engineering science of the KUL are useful. These following courses are used to look in a more mathematical and structured way at the investigated data.

Viewing the fact that an image is represented by an array, Linear Algebra is very important. From basic operations like multiplying matrices up until more difficult acts like convoluting

a matrix, has algebra an important role. Note that each picture taken from the Kinect v2 is viewed as a large matrix with roughly 350000 elements. A knowledge of Numerical Mathematics is advised to get an idea which influence measurement errors have on all of the calculations with this very large matrices. This course, together with Applied Informatics in Python, gives the tools to investigate the time complexity of the algorithm. When this project continues in a further stage, the course of Information Transmission and Processing could be used. Given this knowledge, the efficient use of memory during the algorithm can be monitored.

8 Conclusion

After four weeks of group gatherings, a lot of research has been done and a decent amount of progress was made. The project will be executed by using a Microsoft Kinect version 2. The combination of a depth and RGB sensor at a reasonable cost are some of the decisive arguments over a standard webcam or an industrial camera.

When obtaining the image from the RGB sensor, the three dimensional matrix is turned into a one dimensional matrix and thus into a greyscale image. A Gaussian blur and an edge detection algorithm are used to further reduce this matrix to a binary array where only the edges are highlighted. Meanwhile, using a Sobel-Feldman operator and a threshold filter, the image obtained from the depth sensor will become clean with a clearly visible outline around the differences in height.

Next, the goal is to, after merging the processed data from the RGB and depth sensor, perfect the positioning of the edges and fill in possible blank spots. Subsequently, a step towards counting the objects will be made. If the amount of progress made per day will stay consistent, all the deadlines should be accomplished in the given timeframe. But perseverance and a critical point of view are needed to successfully finish this assignment.

9 References

iiiiiii HEAD =====Dirac, P. A. M. (1981). *The principles of quantum mechanics*. Clarendon Press. *Greyscale*. (n.d.). Retrieved from <https://en.wikipedia.org/wiki/Grayscale> Jiao, J., Yuan, L., Tang, W., & Wu, Q. (Nov 2017). *Kinect v1 and kinect v2 fields of view compared*. Retrieved from https://www.researchgate.net/figure/Kinect-v2-sensor-working-on-a-photographic-tripod_fig1321 Kernel (image processing). (n.d.). Retrieved from [https://en.wikipedia.org/wiki/Kernel\(image_processing\)](https://en.wikipedia.org/wiki/Kernel(image_processing)) Mathworks. (2018). Retrieved from <https://nl.mathworks.com/help/> R. Fisher, A. W., S. Perkins, & Wolfart, E. (n.d.-a). *Laplacian of gaussian*. Retrieved from <https://homepages.inf.ed.ac.uk/rbf/HIPR2/log.html> R. Fisher, A. W., S. Perkins, & Wolfart, E. (n.d.-b). *Mean filter*. Retrieved from <https://homepages.inf.ed.ac.uk/rbf/HIPR2/mean.html> Smeenk, R. (11 Mar 2014). *Kinect v1 and kinect v2 fields of view compared*. Retrieved from <http://smeenk.com/kinect-field-of-view-comparison/> Sobel, I. (2014, 02). An isotropic 3x3 image gradient operator. *Stackoverflow*. (2018). Retrieved from <https://stackoverflow.com/questions/2166326a9c29d0e91c0a8d7a66c9c3672b8bfd4>

10 Appendix

A: Matlab Code RGB sensor

```
1 clearvars
2 img = imread('kinect/foto RGB 3.png'); % Load picture
3 A = greyscale(img); % Convert image to grayscale
4 A = symImgCrop(A, 50); % Crop image so it's the same size.
5 A = gaussian_blur(mean_blur(A)); % Filters
6
7 %% Method 1: First greyscale, then blur, then threshold, then edge
  detection.
8 B = threshold(A); % Threshold image
9 C = ~edge2_detect(B, 3);
10 D = remove_boundary(C, 25);
11 subplot(2,2,1), imshow(A, []);
12 title("Input (after blur)");
13 subplot(2,2,2), imshow(B, []);
14 title("After thresholding");
15 subplot(2,2,3), imshow(D, []);
16 title("After edge detection & boundary removed");
17
18 %% Method 2: First greyscale, then trheshold with background, than
  edge detection
19 bg = imread('kinect/foto RGB 1.png'); % Load background image
20 bg = greyscale(bg); % Convert image to grayscale
21 bg = symImgCrop(bg, 50); % CROP IMAGE SO IT's the same size.
22 bg = gaussian_blur(mean_blur(bg)); % Filters
23
24 B = threshold_lvm_background(A, bg); % Threshold with background
25 C = ~edge2_detect(B, 3); % Detect edges.
26 D = remove_boundary(C, 25); % Remove boundary around image.
27 subplot(2,2,1), imshow(A, []);
28 title("Input (after blur)");
29 subplot(2,2,2), imshow(B, []);
30 title("After thresholding");
31 subplot(2,2,3), imshow(D, []);
32 title("After edge detection & boundary removed");
33
34 %% Method 3: First greyscale, then blur, then edge detect then
  threshold and then noise removal
35 first_edge_detect = edge_detect(A); % Laplacian edge detection
36 without_noise_removal = threshold_edge(remove_boundary(
    first_edge_detect, 15)); % Remove boundary around image &
    threshold the edges.
37 with_noise_removal = noise_deletion(without_noise_removal, 3); %
```

```

    Noise removal
38 subplot(2,2,1), imshow(A, []);
39 title("Input (after blur)");
40 subplot(2,2,2), imshow(first_edge_detect, []);
41 title("After edge detection");
42 subplot(2,2,3), imshow(without_noise_removal, []);
43 title("Threshold without noise removal");
44 subplot(2,2,4), imshow(with_noise_removal, []);
45 title("Method 2 with gaussian and mean blur");
46
47
48 function result = threshold_ivm_background(img, bg)
49     % DIMENSIONS MUST MATCH
50     % Compare pixel at img(row, col) with bg(row, col).
51     % if bg(row, col) - D <= img(row, col) <= bg(row, col) + D
52     %     The pixels are defined as background!! (= white)
53
54     D = 10;
55     WHITE = 1;
56     BLACK = 0;
57
58     matrix_size = size(img);
59     MAXROW = matrix_size(1);
60     MAXCOLUMN = matrix_size(2);
61
62     result = zeros(MAXROW,MAXCOLUMN,1);
63     for row=1:MAXROW
64         for col=1:MAXCOLUMN
65             if img(row, col) <= bg(row, col) + D && img(row, col) >=
                bg(row, col) - D
66                 % Classified as background
67                 result(row, col) = WHITE;
68             else
69                 % Not background
70                 result(row, col) = BLACK;
71             end
72         end
73     end
74
75 end
76
77 function cropped_img = symImgCrop(img, cutted_edge_size)
78     original_img_size = size(img);
79     original_max_row = original_img_size(1);
80     original_max_column = original_img_size(2);
81

```

```

82     cropped_img = zeros(original_max_row - 2*cuttedge_size ,
83                           original_max_column - 2*cuttedge_size ,1);
84
85     for row=cuttedge_size:original_max_row - cuttedge_size
86         for col=cuttedge_size:original_max_column -
87             cuttedge_size
88             cropped_img(row - cuttedge_size + 1,col -
89                 cuttedge_size + 1) = img(row,col);
90         end
91     end
92 end
93
94 function nes = noise_deletion(img,window)
95     matrix_size = size(img);
96     MAXROW = matrix_size(1);
97     MAXCOLUMN = matrix_size(2);
98     side = floor(window/2);
99     nes = img;
100
101     for col=side+1:MAXCOLUMN-side
102         for row=side+1:MAXROW-side
103             list=zeros(window);
104             q=1;
105             for i=-side:side
106                 for j=-side:side
107                     list(q) = img(row+i , col+j);
108                     q = q+1;
109                 end
110             end
111             list=sort(list);
112             nes(row,col) = list(floor((window^2)/2)+1);
113         end
114     end
115 end
116
117 function result = remove_boundary(img, remove_size)
118     matrix_size = size(img);
119     MAXROW = matrix_size(1);
120     MAXCOLUMN = matrix_size(2);
121
122     result = zeros(MAXROW,MAXCOLUMN,1);
123     for row=1:MAXROW
124         for col=1:MAXCOLUMN
125             if row < remove_size || col < remove_size || row > (
126                 MAXROW - remove_size) || col > (MAXCOLUMN -
127                     remove_size)

```

```

123         % Inside boundary ==> needs to be white (= 1)
124         result(row, col) = 1;
125     else
126         result(row, col) = img(row, col);
127     end
128
129     end
130 end
131
132
133 function thresholded_img = threshold_edge(img)
134     threshold_value = 2;
135     %most_occurring = mode(img) + 100;
136     %threshold_value = most_occurring(1);
137
138     matrix_size = size(img);
139     MAXROW = matrix_size(1);
140     MAXCOLUMN = matrix_size(2);
141     THICKNESS = 3;
142
143     thresholded_img = zeros(MAXROW, MAXCOLUMN, 1);
144     for row=1:MAXROW
145         for col=1:MAXCOLUMN
146             if img(row, col) > threshold_value
147                 value = 1;
148                 for i=1:THICKNESS
149                     % Create thicker edges (edges of THICKNESS
150                     % pixels thick)
151                     if (col - i) > 0
152                         thresholded_img(row, col-i) = 0;
153                     end
154                 end
155             else
156                 value = 0;
157             end
158             thresholded_img(row, col) = value;
159         end
160     end
161
162 function mean_blurred = mean_blur(img)
163     mean = (1/9) * [ 1 1 1; 1 1 1; 1 1 1];
164     mean_blurred = conv2(img, mean);
165 end
166
167 function gaussian_blurred = gaussian_blur(img)

```

```

168     gaussian = (1/159) * [2 4 5 4 2; 4 9 12 9 4; 5 12 15 12 5; 4 9
169         12 9 4; 2 4 5 4 2;];
170     gaussian_blurred = conv2(img, gaussian);
171 end
172 function edge2 = edge2_detect(img,intolerance)
173     matrix_size = size(img);
174     MAXROW = matrix_size(1);
175     MAXCOLUMN = matrix_size(2);
176     edge2 = zeros(MAXROW,MAXCOLUMN,1);
177     THICKNESS = 2;
178
179     % Horizontaal laten checken voor edges.
180     previous_value = img(1,1);
181     for row=1:MAXROW % We gaan elke rij af
182         for col=1:MAXCOLUMN
183             i=1;
184             flag = 0;
185             if img(row, col) == 1 && previous_value == 0
186                 % DUS: Het begin van een object. (hele tijd wit, nu
                    zwart), flag voor intolerantie controle
                    aanzetten.
187                 flag = 1;
188             elseif img(row, col) == 0 && previous_value == 1
189                 % DUS: Het einde van een object (hele tijd zwart,
                    nu wit), flag voor intolerantie controle
                    aanzetten.
190                 flag = 1;
191             end
192
193             %%Intolerantie controle
194             while i <= intolerance && flag && col+i <= MAXCOLUMN
195                 if img(row,col-1+i) ~= img(row,col+i)
196                     flag = 0;
197                 end
198                 i=i+1;
199             end
200
201             % Eertse maal edgematrix vullen
202             if flag
203                 edge2(row, col) = 1;
204
205                 for i=1:THICKNESS
206                     % Create thicker edges (edges of THICKNESS
                        pixels thick)
207                     if (col - i) > 0

```

```

208         edge2(row, col-i) = 1;
209     end
210 end
211 else
212     edge2(row, col) = 0;
213 end
214
215     previous_value = img(row, col);
216 end
217
218 previous_value = img(row,1);
219 end
220
221 % Verticaal controleren op edges.
222 previous_value = img(1,1);
223 for col=1:MAXCOLUMN % We gaan elke kolom af
224     for row=1:MAXROW
225         i=1;
226         flag = 0;
227         if img(row, col) == 1 && previous_value == 0
228             % DUS: Het begin van een object. (hele tijd wit, nu
                zwart), flag voor intolerantie controle
                aanzetten.
229             %value = 1;
230             flag = 1;
231         elseif img(row, col) == 0 && previous_value == 1
232             %DUS: Het einde van een object (hele tijd zwart,
                nu wit), flag voor intolerantie controle
                aanzetten.
233             %value = 1;
234             flag = 1;
235         end
236
237         % Intolerantie controle
238         while i <= intolerance && flag && row+i <= MAXROW
239             if img(row-1+i, col) ~= img(row+i, col)
240                 flag = 0;
241             end
242             i=i+1;
243         end
244
245         % Enkel nullen overriden
246         if flag
247             edge2(row, col) = 1;
248             for i=1:THICKNESS
249                 % Create thicker edges (edges of THICKNESS

```

```

250         pixels thick)
251         if (row - i) > 0
252             edge2(row - i, col) = 1;
253         end
254     end
255
256     previous_value = img(row, col);
257 end
258
259 previous_value = img(1, col);
260 end
261
262 end
263
264 function edge = edge_detect(img)
265     klaplace=[0 -1 0; -1 4 -1; 0 -1 0]; % Laplacian
266     filter kernel % convolve
267     edge=conv2(img, klaplace);
268     test img with
269 end
270
271 function thresholded_img = threshold(img)
272     threshold_value = 125;
273     %most_occurring =mode(img) +100;
274     %threshold_value = most_occurring(1);
275
276     matrix_size = size(img);
277     MAXROW = matrix_size(1);
278     MAXCOLUMN = matrix_size(2);
279
280     thresholded_img = zeros(MAXROW,MAXCOLUMN,1);
281     for row=1:MAXROW
282         for col=1:MAXCOLUMN
283             if img(row, col) > threshold_value
284                 value = 1;
285             else
286                 value = 0;
287             end
288             thresholded_img(row, col) = value;
289         end
290     end
291
292 end
293
294 function grey = greyscale(img)

```

```

293     matrix_size = size(img);
294     MAXROW = matrix_size(1);
295     MAXCOLUMN = matrix_size(2);
296
297     grey = zeros(MAXROW,MAXCOLUMN,1);
298     for row=1:MAXROW
299         for col=1:MAXCOLUMN
300             R = img(row, col, 1);
301             G = img(row, col, 2);
302             B = img(row, col, 3);
303             grey(row, col) = 0.2989 * R + 0.5870 * G + 0.1140 * B ;
304             %These are two methods for grayscaling.
305             %grey(row, col) = (R + G + B)/3;
306         end
307     end
308 end

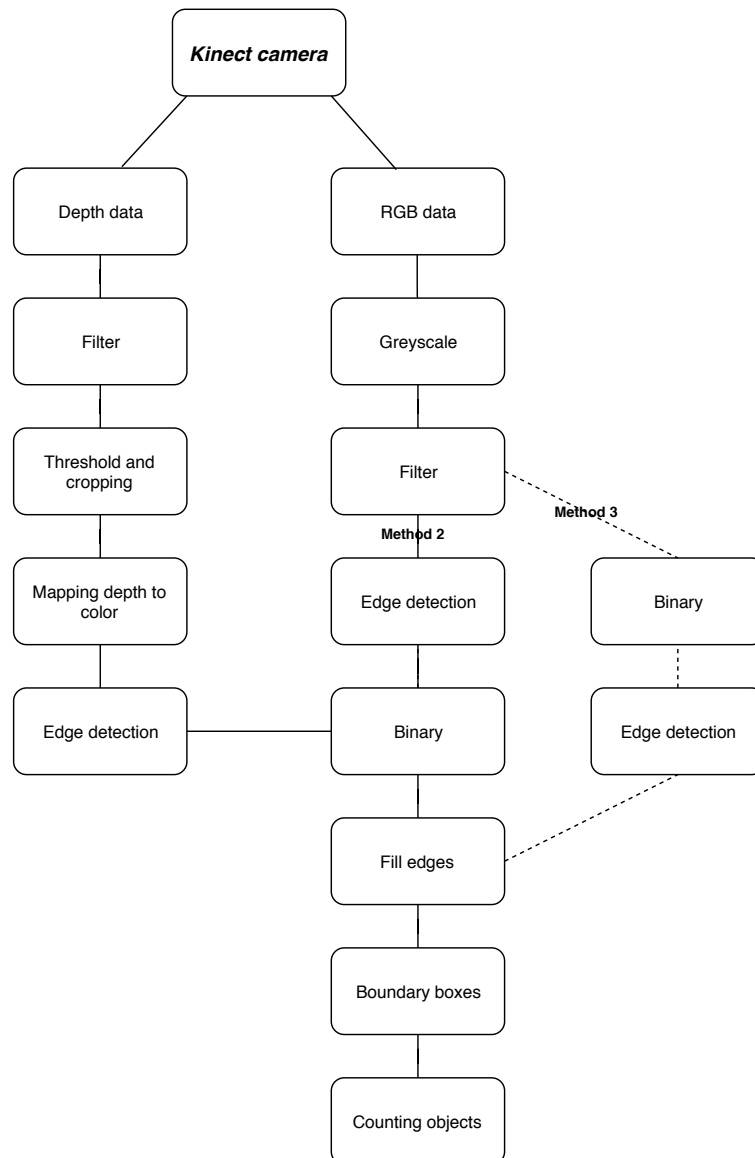
```


B: Matlab Code depth sensor

```
1 %processing the image using the depthsensor
2
3
4 %treshhold values
5 min_tresh = 30;
6 max_tresh = 500;
7
8 % get image from depth sensor
9 depth = getsnapshot(depthVid);
10
11 %run the sobel operator
12 shapes = sobel_operator(depth);
13
14 %run the treshhold filter
15 shapes = treshhold(shapes, min_tresh, max_tresh);
16
17 %look at the result
18 image(depth);
19
20 function shapes = sobel_operator(img)
21
22     X = img;
23     Gx = [1 +2 +1; 0 0 0; -1 -2 -1]; Gy = Gx';
24     temp_x = conv2(X, Gx, 'same');
25     temp_y = conv2(X, Gy, 'same');
26     shapes = sqrt(temp_x.^2 + temp_y.^2);
27 end
28
29 function treshholded = treshhold(img, min_tresh, max_tresh)
30
31     matrix_size = size(img);
32
33     MAXROW = matrix_size(1);
34
35     MAXCOLUMN = matrix_size(2);
36
37     for row = 1 : MAXROW
38         for col = 1: MAXCOLUMN
39             if (img(row, col) > min_tresh) && (img(row, col) <
                max_tresh)
40                 img(row, col) = 1;
41             else
42                 img(row, col) = 0;
43             end
44         end
45     end
46 end
```

```
44         end
45     end
46     tresholderd = img;
47 end
```

C: Diagram



D: Gantt chart

teamgantt

Created with Free Edition

