Analysis of well-being of pigs by means of video sequences

Ricard Moya Mata

Summary (English)

The Danmarks Tekniske Universitet (DTU) collaborates with the Danish Meat Research Institute in order to implement an automated method which takes care of the well-being of pigs on the slaughterhouse. The method will try to survey the welfare of pigs during the process from when they arrive at the slaughterhouse, they are unloaded and finally they go to the CO^2 chamber.

By means of video sequences, the task has consisted in implementing an algorithm that obtains the maximum amount of information from the pigs on the slaughterhouse. Afterwards, the information about them is stored and used as a tool for extracting features about the abnormal behaviour of the pigs.

Summary (Catalan)

La Danmarks Tekniske Universitet (DTU) col·labora amb el Danish Meat Research Institute (DMRI) per a implementar un mètode automatitzat que s'ocupa de la inspecció del benestar dels animals a l'escorxador. El mètode supervisa el seu comportament des que arriben amb camió a l'escorxador, són descarregats i finalment enviats a la càmara de CO^2.

La tasca ha consistit en la implementació d'un algorisme que obté el màxim de informació mitjançant l'us de seqüències de vídeo, que provenen de les càmeres de video-vigilància dels porcs dins de l'escorxador. Seguidament, la informació obtinguda és emmagatzemada i usada com a eina per a la detecció de comportaments anormals del bestiar.

Summary (Spanish)

La Danmarks Tekniske Universitet (DTU) colabora con el Danish Meat Research Institute (DMRI) para implementar un método automatizado que se ocupa de la inspección del bienestar de los animales en el matadero. El método inspecciona su comportamiento desde que llegan en camión al matadero, son descargados y, finalmente, enviados a la cámara de CO^2.

La tarea consistió en la implementación de un algoritmo que obtiene el máximo de información mediante el uso de secuencias de vídeo, que proceden de las cámaras de video vigilancia de los cerdos dentro del matadero. Seguidamente, la información obtenida es almacenada y usada como herramienta para la detección de comportamientos anormales del ganado.

Contents

List of Figures

Introduction

The increasing demand of the food industry to improve the quality in their products has lead the investigation opportunities in the field of the meat industry. Moreover, we have to bear on mind that the animal welfare is more on focus nowadays, and handling of live animals before slaughter is a topic gathering constant attention and discussion. The Danish Meat Research Institute has decided to combine these facts to start a research in this area.

Taking animal welfare into account is lucrative. Primarily, the quality of the meat from animals that are well treated is better, and also gentle treatment results in fewer injuries to the animals, entailing to the survival of more animals until the slaughter time. Legislation in the EU and many other parts of the world is very focused on the handling of living animals. Moreover, interest groups and consumers are very aware of animal welfare.

The DMRI[1] has large competences within animal welfare. They develop cost-effective solutions to improve animal welfare during transport, lairage and stunning, and they assist in implementing procedures and systems that contribute to proper animal welfare at the day of slaughter. Furthermore, they develop systems and methods to control and document the level of animal welfare towards customers and authorities.

---

[1] Danish Meat Research Institute

Some of the examples of the DMRI interception on the meat industry to improve the quality of the slaughtering facilities are:

Best practice for building vehicles for animal transport: [1] Aforementioned, the DMRI has the competences and the tools to ensure animal welfare from the pigs and cattle are delivered, from the farmers to the slaughter. For example, they have elaborated handbooks (HST[2]/HKT[3] handbooks) about the construction of vehicles for transport of pigs and cattle. When the directions of the books are followed, the best conditions are present for transporting animals with a minimum of injuries.

Optimal construction of receipt and stunning are almost at no extra cost free of charge: [1] A carefully prepared construction of loading ramp, lairage and driveways all the way to stunning at the slaughterhouse is crucial to minimize the stress that the animals experience when they are moved around in an unfamiliar environment. Often, when the unloading facilities are being examined, only a small investment can lead to notable improvement of animal welfare, more efficient driving to stunning and an improvement in meat quality.

As it can be seen, the DMRI focuses all its effort on evaluating and improving animal welfare in all the tasks carried out during transportation, off-loading, lairage, driveway, stunning and slaughtering. Therefore, the evaluation of the animal welfare becomes a very important job: the new methods to improve animal well-being during the slaughtering process will rely on it.

The DMRI has taken a step forward in the process and has decided to automatize the process of taking care of the animal welfare. The method will try to survey the behaviour of pigs during the process from when they arrive at the slaughterhouse, when they are unloaded and finally they go to the CO^2 chamber.

---

[2]Handbook Sweine Transport
[3]Handbook Cattle Transport

1.1   Main Goal

The aim of this project is to automatize the recognition of the well-being of animals thanks to the clues that the vets of the DMRI have provided us. The main concept is to create a method for abnormal pig behaviour recognition in video recordings.

In the first step, pigs' movement trajectories should be extracted from video recording. In the second step, using unsupervised or semi-supervised methods to build a quality control system for abnormal pig behaviour detection. Some of the abnormal behaviour indicators are already known. For instance, it is normal to find a pig lying or standing on the ground, while having the pig sitting is a symptom of abnormality. Pigs biting each other is another example of abnormal behaviour. Phd student Pia Brant at DMRI is analysing pigs from the biological point of view and gives some explanations to some of the abnormal behaviours. The interest of the project is focused on discovering more of these indicators.

In my particular case, my task consists in obtaining the maximum amount of information from the pigs on the slaughterhouse by means of video sequences given by the DMRI. This project tries to obtain as much information as possible regarding the behaviour of pigs in the corridors of a slaughterhouse.

The clues that the vets have given us about the animal well-fare point out to a complete tracking of the pigs as the best solution to achieve our aim. All we can expect to obtain from the videos are the position, the direction of the movement, the orientation and the size of each pig in all frames. Moreover, a human eye cannot extract more information of the videos we have been given. Then, the output of the algorithm provides the information to detect the signs of stress that the whole method is looking for.

The Tracking Algorithm

The object tracking is one of the most challenging tasks on the computer vision. There are plenty of disparate techniques that are able to track all kind of different object in such all kind of situations. The key of the question lies on the choice of the optimal tracker. Therefore, the analysis of the given images becomes something crucial for a successful project (look at section 2.0.1).

The goal of a multiple object tracking algorithm is to estimate the trajectory of the interesting moving objects in the image while they are moving around the scene.

Every video analysis is based in three key points: detection, tracking and analysis. First of all, the interesting moving objects have to be detected and isolated. Once this task has been done, it is time to track the aforementioned objects through all the image sequences. Finally, the analysis of the obtained outputs will show the results of the method: such as orientation, centre, area and edges of certain objects.

Most of the algorithms start from the premise that some (high level) information known before the tracking that we can introduce on the algorithm. In consequence, the algorithm imposes constraints on the motion, the size and the appearance of the objects. For instance, in this particular case it can be assumed that the pigs motion can not be more than a certain displacement in the

image. This can be integrated on the algorithm and facilitate the tracking task.

2.0.1   Analysis of the given images

According to these premises, it has been built a tracking algorithm that best fits into the particular case we have been given. There are some complex troubles that make the tracking task difficult. Watching carefully the images we can enumerate the most distinguished characteristics that will make the tracking an elaborated task:

1. Noise: bad quality of the images.

2. Object occlusions

3. Projection from the real 3D to the 2D image.

4. Complicated object shapes

5. Illumination changes of the scene

Looking at the Figure 2.1 I proceed to analyse the input video for the tracking algorithm:

- **Image quality:** the video sequences are provided by a security camera located on the top of the slaughterhouse. As it is usual with this kind of cameras, the output images have not a good quality of image. The low resolution of the images and the lack of colour on them increase the difficulty of the method.

- **Location of the camera:** the camera is focusing on the animals with approximately a 45 degree angle to the floor. That makes a huge difference of the size of the pigs depending on their position in the image.

- **Scene:** the fixed camera and the restricted area where the pigs are able to move makes a good point for the tracking algorithm. The illumination is constant with no significant changes, the contrast of the objects to track is remarkable and their shape is quite similar during the whole sequence.

After the review of the images it is time to start building the algorithm. The deepest knowledge on the specific features of the video images makes the choice of the tracking technique easier.

(a) frame 10

(b) frame 29

(c) frame 53

(d) frame 78

(e) frame 101

(f) frame 125

(g) frame 177

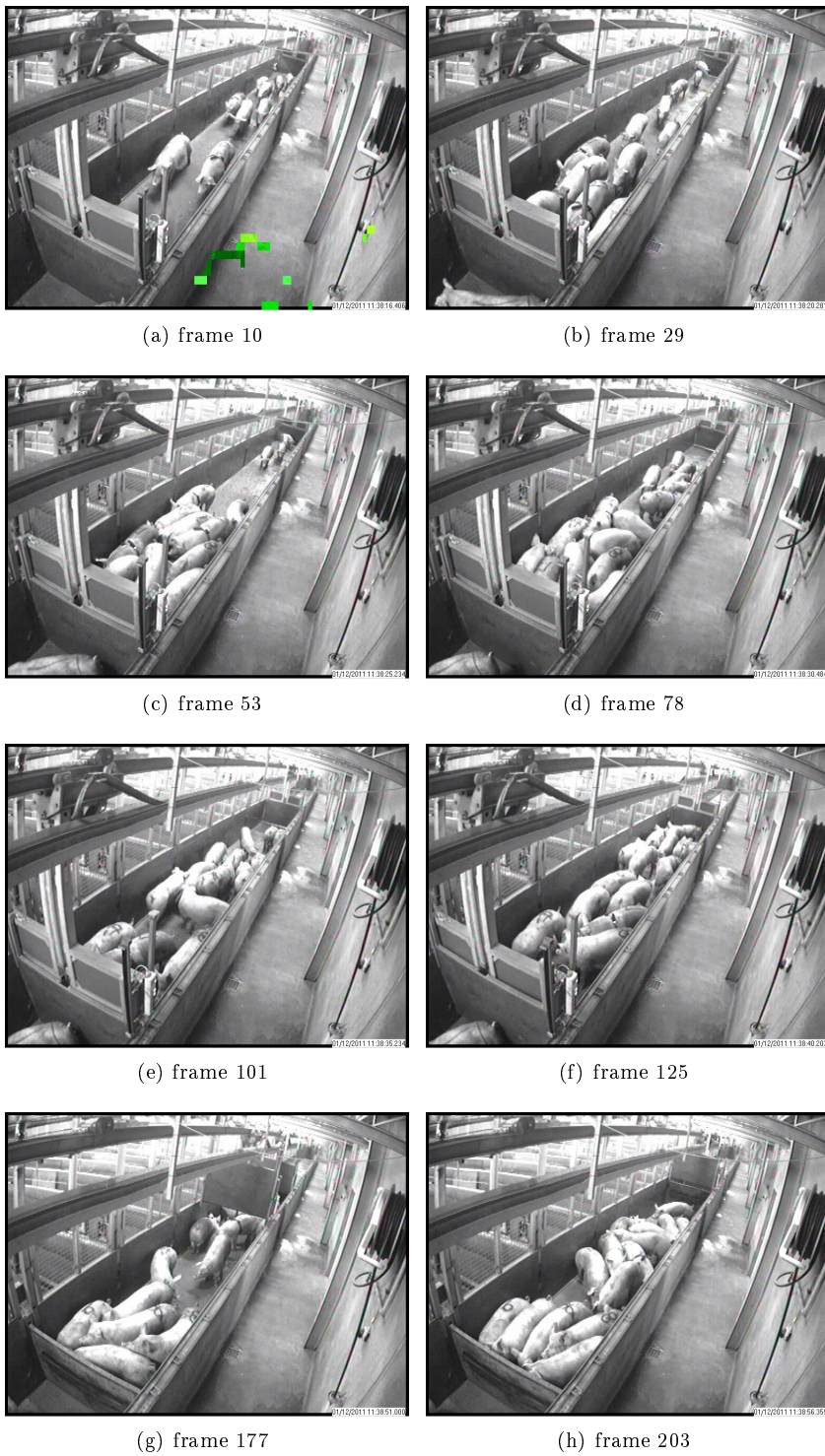(h) frame 203

Figure 2.1: Used video sequence to implement the tracking algorithm

During the first weeks of the project, a review of some of the related investigations was done. Some of them, the similar ones to our project were tested, but neither of them has been a successful solution. Then, the challenge became harder. Finding a good tracking object became the main task during several weeks.

2.1 Tracking features

On this first stage, it has been assumed that the initialization of the location of each pig, and the corresponding localization of the windows fitting on the shape of the pig (object detection algorithm) has been done. This information is taken as the starting point for the features tracking algorithm. Therefore, the problem is reduced to the tracking of the pig inside the located windows. Afterwards, the initialization and windowing problem will be tackled.

After some weeks of reading literature and several tests about the suitable techniques for making the tracking of the objects, the chosen option has been a feature extraction. On the first attempts to find a good algorithm, it was opted for applying the SIFT [6] [5] (Scale-Invariant Feature Transform) which is a very interesting method, but after various tests it was seen that the number of features extracted per pig were not enough to do a successful tracking. Despite that, the tracking of the few extracted features was quite good, then, the key was to find more features, and these are the DAISY features [7].

2.1.1 Descriptor for Dense Wide-Baseline Stereo Matching

This descriptor, called DAISY, is very fast and efficient to compute. It depends on histograms of gradients like SIFT and GLOH but uses a Gaussian weighting and circularly symmetrical kernels. This allows a great speed and efficiency for dense computations. 200-length descriptors can be computed for every pixel in a 800x600 image in less than 5 seconds while the SIFT features take more than 250 seconds.

The Daisy descriptor

As we can see on Figure 2.2, each circle represents a region where the radius is proportional to the standard deviations of the Gaussian kernels and the '+' sign represents the locations where they sample the convolved orientation maps, its center being a pixel location where we compute the descriptor. By overlapping the regions they achieve smooth transitions between the regions and a degree of rotational robustness. The radii of the outer regions are increased to have an equal sampling of the rotational axis which is necessary for robustness against rotation.
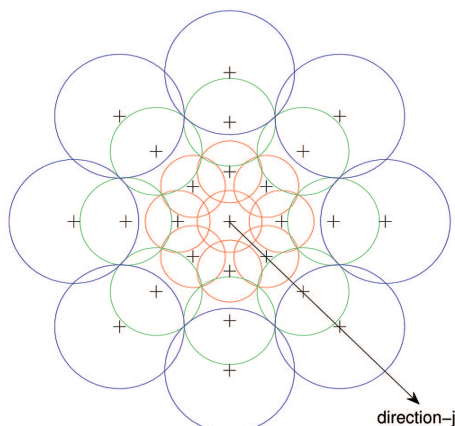
Figure 2.2: Daisy descriptor

For a given input image, it is first computed eight orientation maps, G, one for each quantized direction, where Go(u,v) equals the image gradient location (u,v) for direction o if it is bigger than zero, else it is equal to zero. With this they preserve the polarity of the intensity change. Each orientation map is then convolved several times with Gaussian kernels of different Σ values to obtain convolved orientation maps for different sized regions. This can be done efficiently by computing all these convolutions recursively.

Applying the descriptor to the algorithm

Once it is known with which descriptor we are working on, it has to be applied in an algorithm that allows the tracking of the descripted object. The algorithm computes the DAISY descriptors inside the window of the current frame, then, it finds the matches between the current frame and the consecutive one. Finally, the window has to be updated on the new location that fits better for the new position of the pig. Consequently can see how it works with more detail:

1. The algorithm computes the DAISY descriptors in each pixel inside the window. Each descriptor becomes a vector of one hundred positions that gives information of the related pixel. To lighten the algorithm, it is possible to compute just one descriptor every 5 or 10 pixels, the final results are not affected by this reduction.

2. Once we have the descriptors it is time to find the matches on the next frame. For each obtained descriptor inside the window, the algorithm

looks for the most similar descriptor on the next frame with one constrain: the pig's movement is not bigger than a square of 25x25 pixels. This information can be introduced on the research of the matches to lighten again the algorithm computation.

The similarity is calculated as the root mean square (RMS) of the descriptors of the two consecutive frames. The chosen pixel is the one with a minimum root mean square. See Equation 2.1:

$$RMS = \sum_{i=1}^{100} \sqrt{(v_b(i) - v_a(i))^2} \qquad (2.1)$$

v_a = descriptor of current frame; v_b = descriptor of next frame

The result of this step can be seen on Figure 2.3. The matching vectors start at the pixel of the current frame and ends at the new found position on the next frame. They are plotted on the current frame.

(a) Matching vectors   (b) Matching vectors zoom
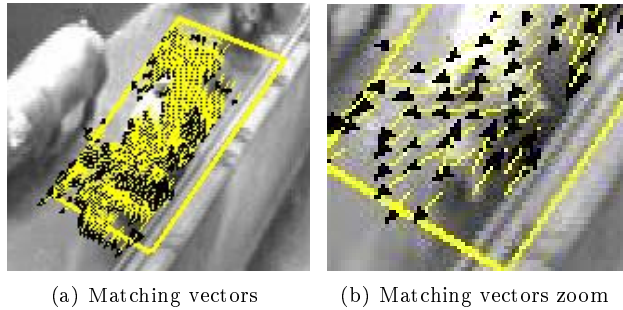
Figure 2.3: To simplify and understand better the images, just a few representation of all the calculated vectors have been plotted.

3. The next step is to update the window to the new location of the pig on the next frame. The process is simple:

   i. The angle of all the matching vectors is computed.
   ii. Then, the mean of these values is extracted to know the mean direction of the window.
   iii. The corresponding matches that have an angle around the mean (variance of pi/15) are selected.
   iv. Finally, the mean module of the selected matches is computed.

   With the mean module and the mean angle the window just has to be updated to the new position on the next frame.

2.1.2   Results of the feature tracking algorithm

As it can be seen on the images on Figure 2.4, the results are quite good and better than expected. Keeping in mind that at the moment there is not any adaptive implementation for the movement of the windows, the algorithm executes perfectly what it has been required. The pictures on Figure 2.4 are an example of that.

The first tests have been done with only one window to simplify the work. It is also interesting to see that the algorithm is also working with multiple windows. As it can be seen on the images on Figure 2.5, it is working quite well. Probably the results are not satisfactory as in the case with just one window, but the algorithm is still working with quite good.

Self-evidently, not everything is perfect and I have detected some errors that I expect to fix with the automatic initialization of the windows and an updating of their localization. On Figure 2.6 I report some of the systematic detected errors.

(a) Frame 13   (b) Frame 46

(c) Frame 61   (d) Frame 89
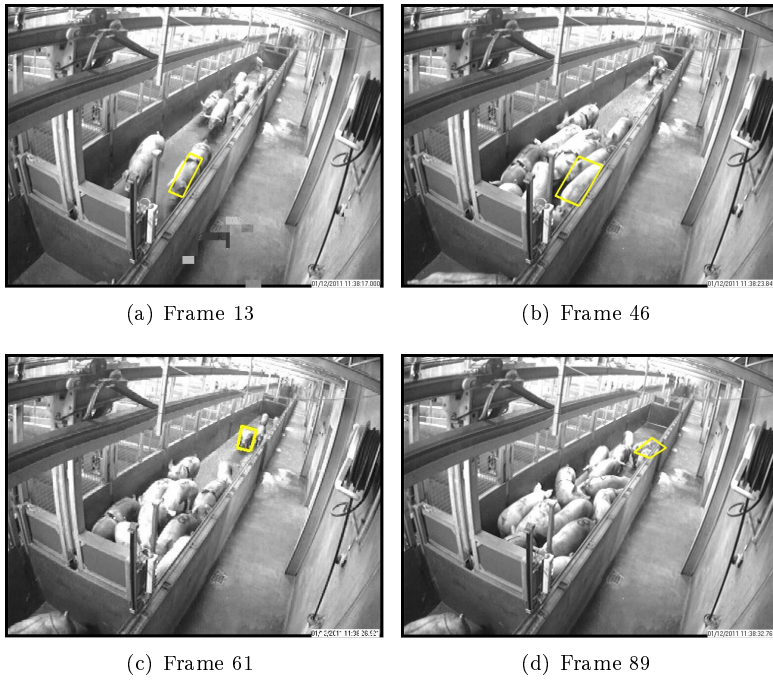
Figure 2.4: Screen shots of the results of the tracking algorithm with one window

(a) Frame 15 3 windows

(b) Frame 59 3 windows

(c) Frame 202 5 windows in static situation

Figure 2.5: Screen shots of the results of the tracking algorithm with multiple windows

(a) Frame 93, pig turning around

(b) Frame 411, 6 windows, tracking error

(c) Frame 47, 9 windows, merge of windows in crowd
situation

Figure 2.6: a) and b) The most common error is the false tracking because of
the influence of the closest pig. c) Another common error is the
false tracking for an influence of the windows around.

2.2 Detection algorithm

As it has been seen on the last section, a part of the tracking algorithm that is able to track successfully a pig inside a located window has been done. This algorithm assumes that has been done an initialization of the location of each pig and the corresponding localization of the windows fitting on the shape of the pig. Then it takes this information as the starting point. Now it is time to build the beginning process. This is clearly a problem of object detection since we have to detect the pigs and to reconstruct a window that approximately takes the shape of a pig.

2.2.1 Object Detection

The main reason of the object detection is to find complementary information about some regions. These regions could signal the presence of interesting objects in the video sequence. Then, this will be used as an automated input for the already implemented tracking algorithm.

As it has been commented on the introduction, the object detection is one of the three key points for any tracking algorithm. The aim of the object detection is to identify the position of the moving object and the size and shape of it on the current frame. Usually, this task is one of the most challenging because of the characteristics of the given video sequences and the moving targets. In order to make the problem easier, some constrains to help finding the animal on the image can be introduced on the algorithm. The number of moving objects, the area to be detected on the video sequence or the kind of objects (animals in this case) will be some of the introduced constrains.

Objective

There are plenty of different techniques of object detection. In my particular case, it is clear that an image blob detector is needed. The purpose of the algorithm is to isolate the animals on every certain amount of frames. Then, using the output information (location, size, orientation,...) to introduce them to the already implemented feature tracking algorithm.

Before going through the algorithm, it is important to understand what is a blob detector. There are some definitions of blobs, but approximately most of them refer to a blob as a region on the image that differ in properties like color

or brightness compared to the surrounding. Others like Lindeberg [3], define a blob as a bright region of interest in a dark background or vice versa, with at least one local extremum. Then, the region is limited by a point where the intensity stops increasing and starts decreasing (for dark blobs or vice versa). Hinz [2] describes the blob just as a rectangle with constant contrast (it becomes a local extremum under scalling).

After this basic knowledge I am able to explain which technique has been used for the object detection and which are the results in this case. The object detection algorithm has been implemented using the laplacian of Gaussian technique. The next paragraphs are dedicated to explain the implementation of the algorithm.

2.2.2   Object detection algorithm. Laplacian of Gaussians.

The laplacian of Gaussians is one of the most common methods used for blob detection. In this case, I am looking for one method that is able to recognize the location and the shape and size of each pig on the image. Given an input image $f(x, y)$, the image is convolved by a Gaussian kernel (equation number 2.2) at a certain scale σ, to give a representation $L(x, y, \sigma) = g(x, y, \sigma) * f(x, y)$. [4]

$$g(x, y, \sigma) = \frac{1}{2\pi\sigma} e^{\frac{(x^2 + y^2)}{2\sigma}} \tag{2.2}$$

Then, the laplacian operator is computed ($\nabla^2 L = L_{xx} + L_{yy}$) and the obtained images have strong positive responses for dark blobs of size $\sqrt{2\sigma}$ and strong negative responses for bright blobs of the same size. As we can see on the next images, we have some examples of the results on Figure 2.7

These results are quite good to perform a correct detection. The only problem resides in the Laplacian operator. It detects edges as well as noise, and that is a problem for the kind of images I have. Even though, it can be easily solved with the application of another related technique, the difference of Gaussians (DoG). The fact is that the Laplacian of the Gaussian operator $\nabla^2 L(x, y, \sigma)$ can also be computed as the limit case of the difference between two Gaussian filtered images.

$$\nabla^2 L(x, y, \sigma) \approx \frac{\sigma}{\Delta\sigma}((L(x, y, \sigma + \Delta\sigma) - (L(x, y, \sigma - \Delta\sigma)) \tag{2.3}$$

Finally, with this approach we obtain satisfactory results that can be seen on Figure 2.8.

(a) Frame 10

(b) Frame 200

(c) Frame 10, $\sigma = 2$

(d) Frame 200, $\sigma = 2$

(e) Frame 10, $\sigma = 8$
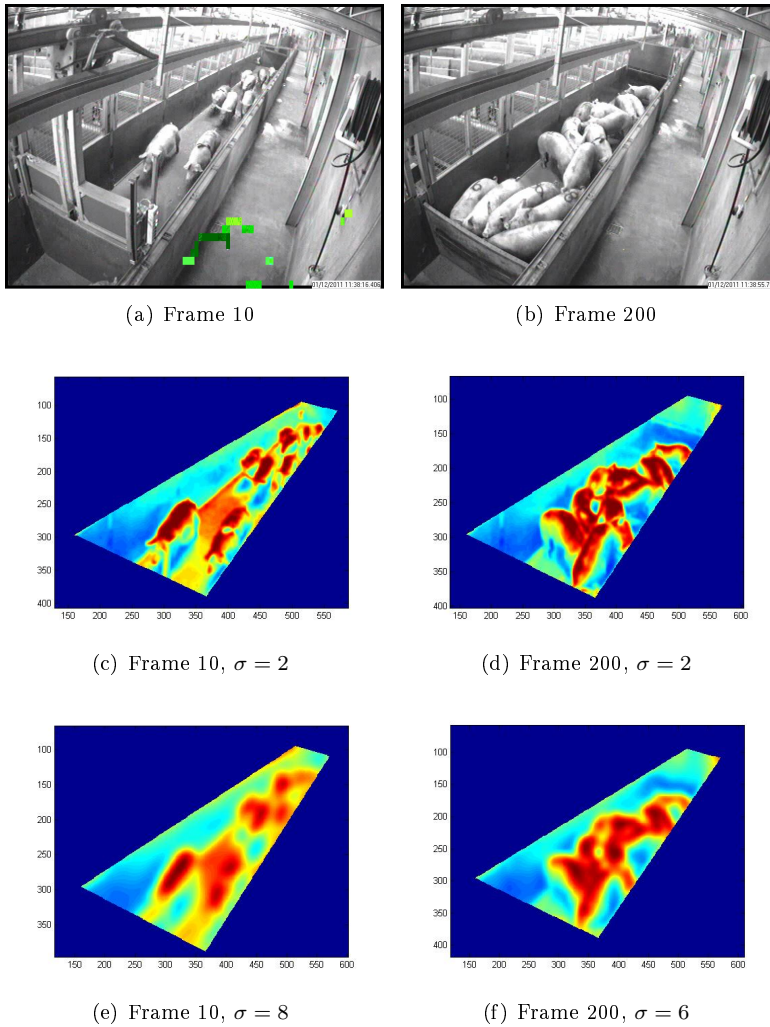
(f) Frame 200, $\sigma = 6$

Figure 2.7: As we can see, the results are quite good. The main problem is that the Laplacian operator detect edges as well as noise.

(a) Frame 10

(b) Frame 200
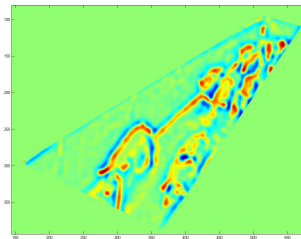
(c) Frame 10, $\sigma_1 = 3, \sigma_2 = 4$
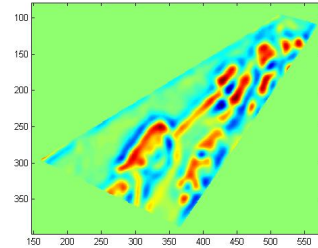
(d) Frame 10, $\sigma_1 = 6, \sigma_2 = 7$
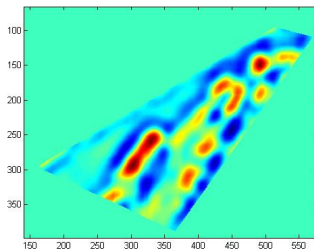
(e) Frame 10, $\sigma_1 = 10, \sigma_2 = 11$
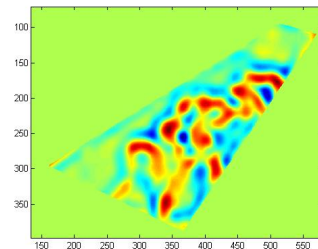
(f) Frame 200, $\sigma_1 = 8, \sigma_2 = 9$

Figure 2.8: The images show a good detection of the location, shape and size of each pig. The problem resides on finding the correct scale.

Once this method has been implemented, there is a scale-base problem for the images. The perspective makes the pigs have a different size in different zones of the images, that means that the correct scale in each zone has to be chosen. To maximize and automatize this problem, an automatic scale selection should be the best option.

At the moment, the object detection with approximate values on the scale of each area, works fine, and on Figure 2.9 can be seen some frames as an example.

2.2.3   Other considerations

When it was started the implementation of this part of the algorithm the objective was to build an object detection algorithm that was just useful to locate and to size the pigs in some frames such as initialization method as an input for the first implemented feature tracking method. After some tests, it was seen that the object detection algorithm exceed the expectations. Actually, when the algorithm is working on an ideal situation (the pigs are isolated and there is no contact within them), it is working almost as a tracking algorithm, giving almost perfect results in some cases.

Hence, after the implementation of both algorithms it has been realized at this point, the challenge is to mix two quite good algorithms (each one on different aspects) and achieve a good and unique tracking algorithm.

2.3   Algorithm assembly

First of all, to carry out the task of assembling the final algorithm it is important
to know the handicaps and the strengths of each algorithm.

	Weak Points	Strong Points
Feature Tracking	Needs an initialization to work. (automatic or manual) • Slow processing time (around 10-20 seconds per frame).	• Robust algorithm against bad conditions. • With a good initialization the results have a good precision.
Object Detection	• Unstable: each frame is like a new frame. The size and the position of the pigs vary too much in every frame. • Working just in perfect conditions: when pigs are isolated and there is no contact within them.	• On ideal conditions works as an almost perfect tracking algorithm. • The processing is 10 times faster than the feature tracking algorithm.

Table 2.1: Strong and weak points of the used methods for the tracking algorithm.

Analysing and comparing both algorithms, it is easy to define a simple route:
The basic running of the algorithm is as it was planned. The object detection
algorithm carries out the initialization part, and gives as an input to the feature tracking algorithm. Then, the feature tracking makes its task and gives
the output of the final tracking algorithm. Now the key resides on taking more
information of the object detection, specially on the cases where it is working
very good. On the next Figure 2.10 we can see the scheme of the algorithm:
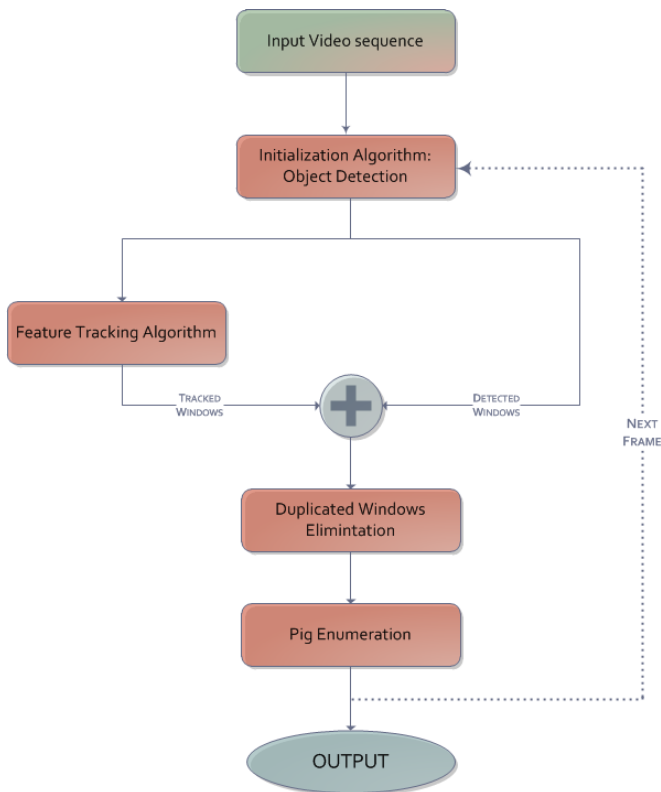
Figure 2.10: Detailed diagram of the complete tracking algorithm

2.3.1   Duplicated Windows Elimination

The main task of this algorithm is to choose the window that fits better on the pig's body. After running both algorithms (feature tracking and object detection), it is usual to have more than one window pointing to the same pig. The algorithm tries to make the best choice keeping just one of the duplicated windows.

As we can observe on the aforementioned table 2.1, the object detection algorithm has some instability and working problems in non-ideal conditions. With this method we try to solve this problem and we try to keep the tracked window (if there is any) and eliminate the detected ones. On Figure 2.11 it can be observed the final results with and without this method:
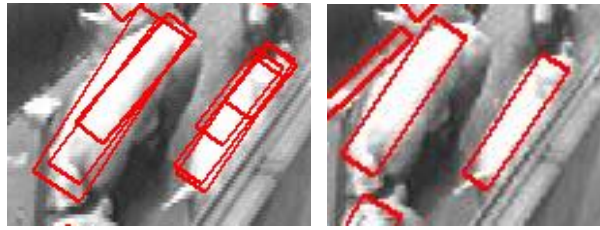
(a) Without the Duplicated Windows Elimination algorithm.

(b) With Duplicated Windows Elimination.

Figure 2.11: Here it can be seen an example of the results of the explained method

The method to eliminate the duplicated windows is a non complex task. The main point is to compare the centres of the final windows and calculate the distances between them. It is easy to see that if some of the centres are closer than the usual width of a window, probably these windows will be pointing the same pig. Giving priority to the feature tracked windows (this method has a better precision pointing the pigs and less instability than the detection), the detected window will be erased. In case of both windows are from the detection method, a random window is erased.

2.3.2 Pig Enumeration

The main task of this method is to make an identification of each pig. This algorithm it is made basically to make easy the job of knowing the trajectories, size, and orientation of each pig on each frame. Then, the vets and the people in charge of the data analysis have better organised information that makes easier to analyse.

The basic functioning of the subroutine is simple: It is easy to know that the pigs has a limited displacement equal to a certain amount of pixels that I decide to call minimum displacement. Then, looking at the final results of the windows on each frame, the windows that are closer than the minimum displacement in the same orientation in consecutive frames will pertain to a same pig. As it can be though, this is not a very precise method to extract this information, but probably is the only possibility with the information that it can be extracted from the images. Furthermore, after some test and manual checking, when the tracking is effective, the pig enumeration is perfect.

(a) Frame 10

(b) Frame 22

(c) Frame 36

Figure 2.9: Here it can be seen some of the final results obtained by the Automatic detection algorithm.

CHAPTER 3

Evaluation and Final Results

The previous chapters have been a revision of the aim of the project, the necessity to build an automated system to identify the stressed animals on the slaughterhouse and finally how to build one of the basic tools of the automatic process. Now it comes the time to evaluate the work I have been doing and the analysis of the final results. It is clear that during the description of the different used methods some partial results have been shown, but in this chapter the final results and the possible applications and uses they can be utilized to will be analysed.

An important part of the work is also evaluating the results. In a tracking algorithm, it is usually easy to see the job of tracking and evaluating the trajectories of the detected objects because of the visual results. Even though, it becomes a time when precision and accuracy take an important part of the algorithm, and it is then when a tool that evaluates how good is your algorithm is very useful to know if your new steps are going forward or on the contrary it is necessary to go back and retake the step before.

3.1   Evaluation

The evaluation method is the tool that defines the accuracy of the elaborated algorithm. There is no rule or theory that defines which is the best way to evaluate the algorithm, but it is clear that it has to be a kind of comparison between the output obtained by the elaborated algorithm and an output that describes the real data the algorithm is trying to calculate. Because of that, it has been necessary to collect the ground truth data manually for the application of the method.

3.1.1   Ground truth data recollection

The toughest task of implementing the evaluation of the output data is the ground truth recollection. It is a manual task consisting on locating each detected object in each frame, as well as annotating all the information that the algorithm is supposed to give as an output. In this case, the four corners of the rectangle that fits the window on the animal, the size of the pigs, the orientation and the centre point of the window have been the variables that have been recollected for each pig on each frame of the video sequence.

The importance of taking a good sample: The recollection of the ground truth data is a hard task that takes long time. We have to bear in mind that for instance, on the given images there are more than 500 frames, with a maximum of 15 pigs in each frame. That means that we have to locate the pig, annotate the size of the pig and the orientation in each frame for each pig, that is to say, 6000 images to analyse.

On the first stage of the evaluation, I have taken a sample of 200 frames, where it can be seen the whole sequence of the pigs entering to the corridor from the door at the back of the images, then they walk all the path until they wait for the next door that brings them to the next station. On the whole video, this sequence is being repeated for groups of 15 pigs.

Finally, the most important part of the ground truth information recollection is the organization of the obtained data. In this case, the output data has been organised in an array of 200 frames, where on each cell (corresponding to each frame) there is another array of structs. Each cell of this second array contains a struct with the information related to each pig on the correspondent frame. We can see a scheme on the Figure 3.1

On the visual section, the chosen representation of the ground truth data is
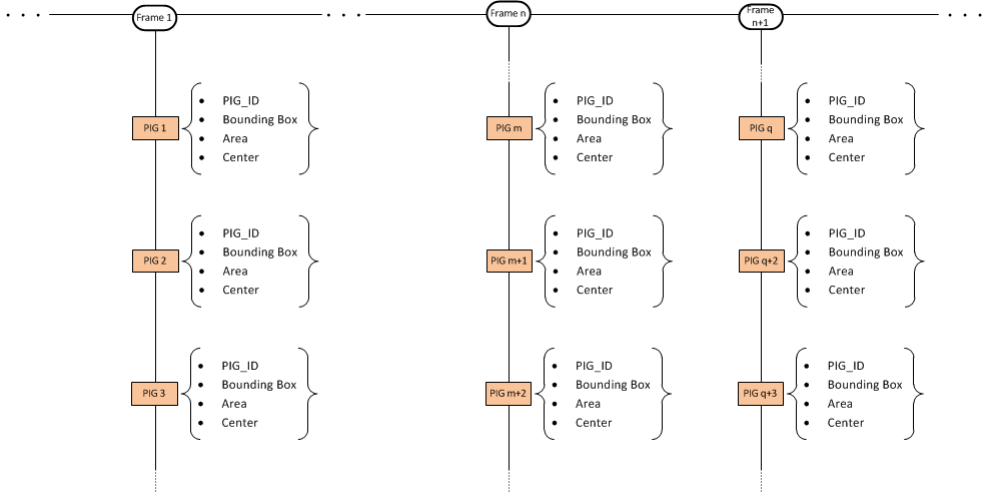
Figure 3.1: Scheme of the the output data organisation.

the same as the visual output of the tracking algorithm. The only represented information is the limits of the bounding boxes, and each pig is represented with a different color. We can see an example on Figure 3.2.

3.1.2 The procedure

Once the ground truth information has been collected, we are able to compare it with the output of the algorithm. Here it is when I bump into some problems that I describe below:

- The algorithm's output does not give an identification for each pig along the frames. That means that there is no way to identify each pig of the ground truth information with the algorithm's output.

- The Bounding Box that defines a pig on the ground truth information is the maximum area that can contain that pig. Then, on the algorithm's output the containing bounding box of each pig is almost always smaller than the ground truth box. That does not mean that the tracking and the detection is wrong, it is just a different way to point to the same pig.

Taking into account these main problems, I have carried out a procedure to obtain an evaluation of the output data.

(a) frame 14

(b) frame 32

(c) frame 59

(d) frame 83

Figure 3.2: Some examples of the graphical ground truth data representation

The procedure: As I have described, the information given by the ground truth data about the pig identification, orientation and trajectory of each pig cannot be used at the moment, because the algorithm is not giving this information correctly enough. Then, I am obligated to use the bounding box information, and I will use the visual information of the algorithm in each frame separately. The steps at this point are:

Frame by frame, I compare the ground truth boxes with the tracked boxes.

1. Once obtained the mask of all the ground truth boxes of the current frame (look at Figure 3.3.a), the first tracked box is compared with the whole ground truth boxes.

2. The difference between the tracked window and the corresponding ground truth window is found. (look at Figure 3.3.b)

3. It is considered that if a tracked window matches more than the 70% with a ground truth window, it is a true positive. If not, it will be a false positive.

4. Finally, the ratio of the true positives over the ground truth windows on each frame give a percentage of the true positives per frame. Then, doing the mean for all the frames the value for all the algorithm is obtained.

$$Evaluation = \frac{true\ positives}{ground\ truth\ positives} \tag{3.1}$$

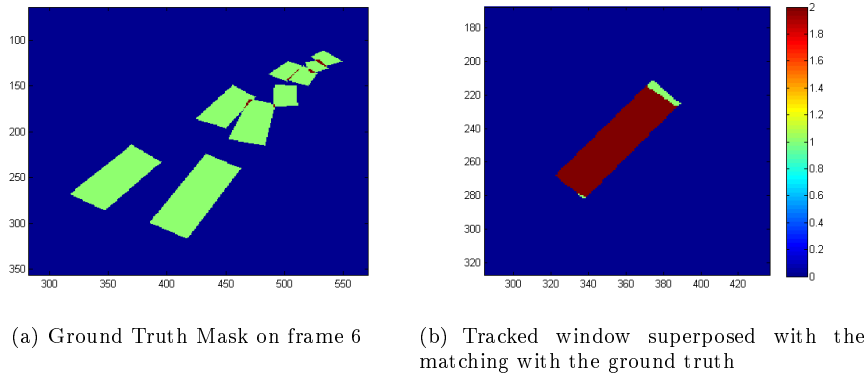(a) Ground Truth Mask on frame 6   (b) Tracked window superposed with the matching with the ground truth

Figure 3.3: (a) Ground truth mask on frame 6. On (b), the yellow pixels represent the area where the tracked window and the ground truth window are not matching.

On the next section, the results of these evaluations will be discussed meanwhile the graphical results are shown.

3.2 Final Results

Until now, the tracking algorithm for pigs on the slaughterhouse has been built. The built tracking algorithm consists on merging the two principal methods for detection and tracking objects. The object detection allows the algorithm to find an image which objects (pigs in this case) have to be tracked for the following method: the feature tracking. After solving some of the problems related to the merging method, it is the time to go through the results.

3.2.1 The results nature

The final output is differentiated by two kinds of information. On the one hand, the output of the algorithm has a numeric nature: as it is explained on section 3.1.1, the information of every output frame is organised in an array of structs, where each cell corresponds to the basic information of each tracked animal. Then, the pig identification number (PIG_ID), the borders of the Bounding Box, the size and the center of each pig per frame can be known. (An example of how organised are they can be seen on Figure 3.1).

On the other hand, the numeric results are hard to understand and very difficult to work with. That is why it is also very important to show the output in a graphical scene. Hence, the best way to show the output is on the same nature of the input: a video sequence.

The graphical results

The graphical results are the visible part of the output of the algorithm. We have to bear in mind that the whole information is on the numerical data saved after the end of the algorithm analysis. But, as humanbeings, the incredulity of everything until we do not see it with our eyes, makes very interesting to represent the results on a video sequence. Furthermore, the graphical representation is a significant help to understand the output results.

First of all, before showing the main results of the algorithm it is important to understand what has been done on the image. The process is simple, and the algorithm just adds more graphical information on the input video sequence.

Building the windows: the result of the object detection algorithm are the

red drawn rectangles made to indicate that there is a tracked pig. As it is explained on the section 2.2, the output of the detection algorithm is a blob that takes the contour of the pig's body. Then, measuring the length (in pixels) of the major and the minor axis, the orientation and the centroid of the ellipse that has the same normalized second central moments as the detected region, is possible to build a rectangle around the detected animal.

The chosen shape for the graphic representation has been a rectangle due to the output of the detection algorithm gives us approximately the trunk of the pig, leaving all the extremities and the head outside the detected part. After some tests, the best shape to include the major part of the extremities and the head without taking parts of other closer animals is the aforementioned rectangle. The final result can be seen on Figure 3.4
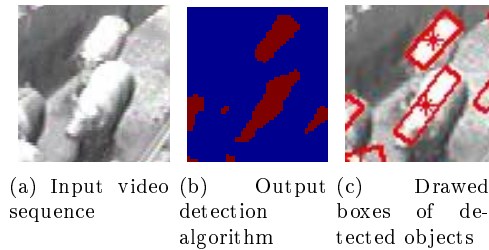

(a) Input video sequence

(b) Output detection algorithm

(c) Drawed boxes of detected objects

Figure 3.4: a)Frame 20 from input video sequence. b)Output of the detection algorithm on frame 20. c) Drawn boxes after detection on frame 20.

CHAPTER 4

Discussion of the final results

Looking back to the beginning of the project, it is important to analyse if the aim of it has been successfully achieved. The main goal of the project is to automatize the process of stress and behaviour recognition of the pigs in the slaughterhouse. Looking into that, the challenge for the image processing was to obtain the maximum of information of the pigs in the slaughterhouse by means of video sequences, that means, knowing the location, size and orientation of each animal in the image every time during the whole sequence.

On the next chapters, concluding the project report, I will focus on the evaluation, discussion and suggestion of possible applications of the information extracted from the given video sequences.

4.1   Discussion of the results

First of all, I will proceed to analyse the obtained results of the first video sequences we were provided. A previous analysis was done before choosing the tracking methods. There, a more detailed analysis can be found , but the main conclusions are:

1. The general quality of the image is very poor. Grayscale, fairly good image resolution and some coding errors in some isolated frames.

2. The location of the camera and the continuous obstructions of the pigs because of the automatic doors.

3. The fixed camera, the constant illumination and the restricted area where the pigs are able to move are positive points for the video tracking.

The expectations for an almost perfect tracking algorithm of the video sequence are not very high, but even though, the effort for having the best tracking on this sequence made and will make easier all the next steps on making suggestions for the technical configuration of the new recordings and having successful results on them.

The whole project has been focused to have usable results for the first images that were given to us. But, as the weeks passed, we realized that the challenge to obtain satisfactory results on this images was more difficult than it was thought at the beginning, and therefore, it would probably be harder to reach the goal I was given.

Despite that, all the efforts have not been in vain, and all this work has helped to have a robust tracking algorithm. Evidence of that are the results of the second video received a few weeks before finishing the investigation. As we will see on the next paragraphs, the results are satisfactory and with an increase on the quality in almost all aspects.

4.1.1   Results for the first video sequence

On the images on Figure 4.1, the final results for the tracking method on the first video sequence are shown. The evaluation of this images gives a promising result:

> The true positives percentage of the output of the algorithm is: **73.2%**

Consequently, that means that the 26.8% of the detections are false detections. Then, the results have to be analysed:

- Visually, all the pigs are treated for at least one window in almost the whole video sequence. Despite the elimination of the window duplicity,

on crowded areas some duplicity of these windows still appear. On the contrary, almost the whole group of animals is always covered by a window, that means that the algorithm is tracking all the pigs during all the video sequence, and that is a strong point.

- On the pig identification (look at section 2.3.2), the non-perfection of the algorithm, the lack of window duplicity and the non-robustness in some areas makes this algorithm unuseful. The pigs identification is implemented assuming a robust tracking of the pigs, and that is why the final results have some errors that do not make optimum their use as a veridical result. Even though, the tests with the ground truth data reveal that the algorithm works perfectly, which makes the pig identification a very reliable algorithm to use on the following jobs.

Concluding the first results, it has to be said that the quality of the obtained results is very high. The high precision demanded by the automated system this project is part of makes the result not useful directly for the system. Even though, the results are very encouraging because the hoped results can be achieved when having better and promising images with a slightly better quality. The proof of that can be found on the next results, that correspond to the input video sequences arrived just a couple of weeks before closing the investigation. Applying and updating the tracking algorithm done for the first images, very high precision and better quality results have been achieved. The next section develops these results.

4.1.2 New video sequence, improving the results

As it has been explained during the report, the task of developing a tracking method with the images we were given was a very challenging task due to the difficulties that were found on the old video sequence. When the investigation was on the testing step of the project plan, we received a new video sequence of the pigs on the slaughterhouse with a new different configuration and with very promising characteristics. On Figure 4.2 some frames of the new received data can be seen.

As it can be observed in a first view, the images have better possibilities on the tracking aspect.

- The most important change on this video sequences are the contrast with the background and the foreground. The pigs are very easy to distinguish from the background due to the strong contrast.

- The size of the pigs is constant. The camera is placed on a fix point, with no inclination with the floor that gives a video sequence without perspective. That eliminates one of the more difficult problems that have to be solved on the anterior video sequence.

- The number of pigs at the same time in the image are much lower than the old sequence and the size of them is much bigger. The algorithm is independent of the number, size and shape of the animals, therefore is not affected by these changes.

- The main challenging task is to distinguish the pigs when they are very close to each other. When there is contact within them the high contrast makes difficult to define the borders of the pigs.

Evaluation of the results

A sample of the most significant frames in terms of the results in the output of the algorithm are shown on Figure 4.3

The tracking algorithm used to obtain this output data is basically the same that has been used for the first results. Due to the facilities imposed by the new data the algorithm has been simplified. In fact, in this case the most general version of the algorithm can be used with almost any special treatment for the images.

Here there are three interesting points to mention about the algorithm:

- **The only method used to track the animals is the object detection:** The feature tracking method is less effective in this case. The pixels inside the animals are very similar, but the matching on the next frames is not successful.

- **The algorithm is 10 times faster:** The fact of the elimination of the feature tracking on the algorithm makes it 10 times faster than the previous one.

- **Elimination of the duplicity of windows:** The elimination of the feature tracking makes disappear the duplicity of windows, one of the most annoying problems on the previous method.

The true positives percentage of the output of the algorithm is: **90.21%**

Precision and accuracy: These are the principal characteristics of the results for the new data. The algorithm turns to a most robustness method and it is precise on the detection of all the movement of the isolated pigs. If there is no contact within the animals, the algorithm works with a 100% of effectiveness. Consequently, the algorithm works perfectly more than the 90% of the time because the conditions are ideal during at least this amount of time.

However, the analysis of the 10% of the time when the algorithm starts having some troubles cannot be skipped. On Figure 4.4 some of the most common problems found on the output sequence can be seen .

1. **Crowd problem:** Sometimes, the pigs pass in crowded groups. This is one of the most challenging tasks. As it can be seen on 4.4.a and 4.4.b, the pig identification is done, but the problem lies in the size detection.

2. **Double identification:** Sometimes the pigs are identified with two different windows. (look at 4.4.c) It is a light problem because both of them are correctly identified but obviously, the size is not correctly found. That can also give some problems on the pig identification.

3. **Double pig inside one window:** Sometimes, when the pigs are close to each other, the algorithm can consider them as an only pig. Here, the solution is easy since the size of the window is the double of one pig window. Then, the algorithm can easily identify that inside that window there are two pigs.

The solutions for these problems do not have a strong complexity. But, due to lack of time, any solution could not be implemented for this small problems.
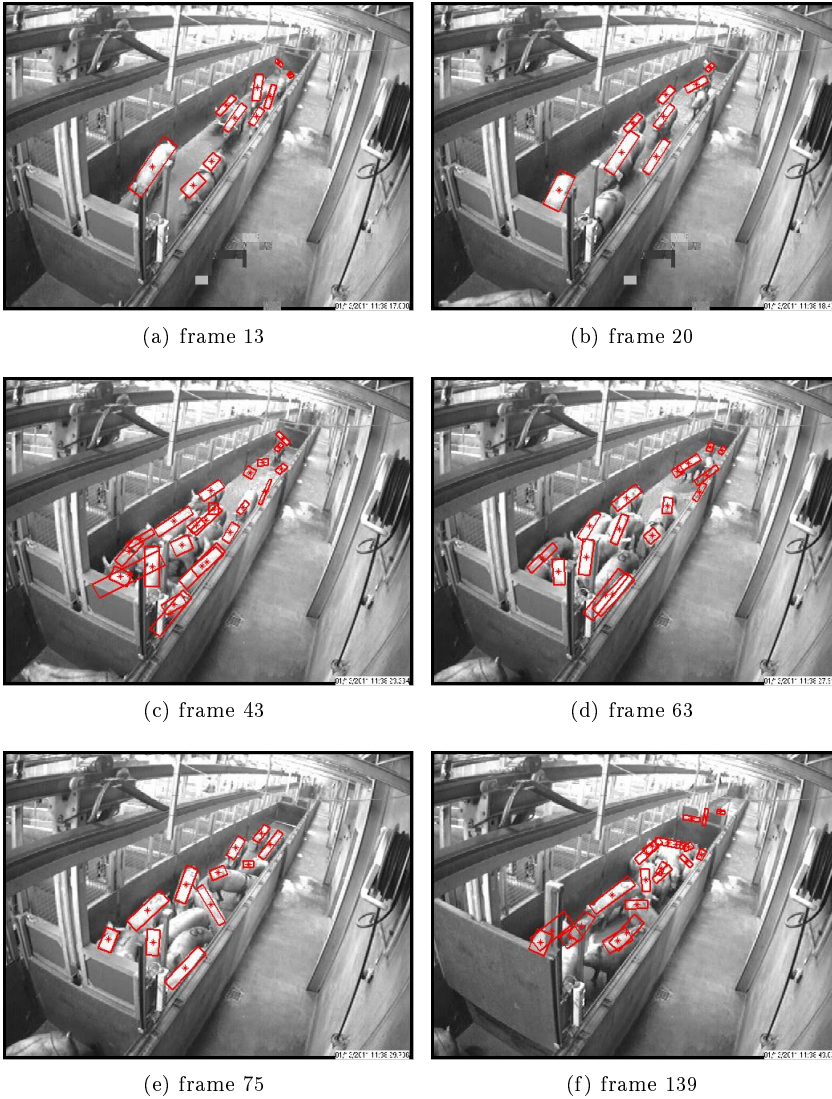
(a) frame 13   (b) frame 20

(c) frame 43   (d) frame 63

(e) frame 75   (f) frame 139

Figure 4.1: Some frames of the final results on the first video sequence

(a) frame 77

(b) frame 148
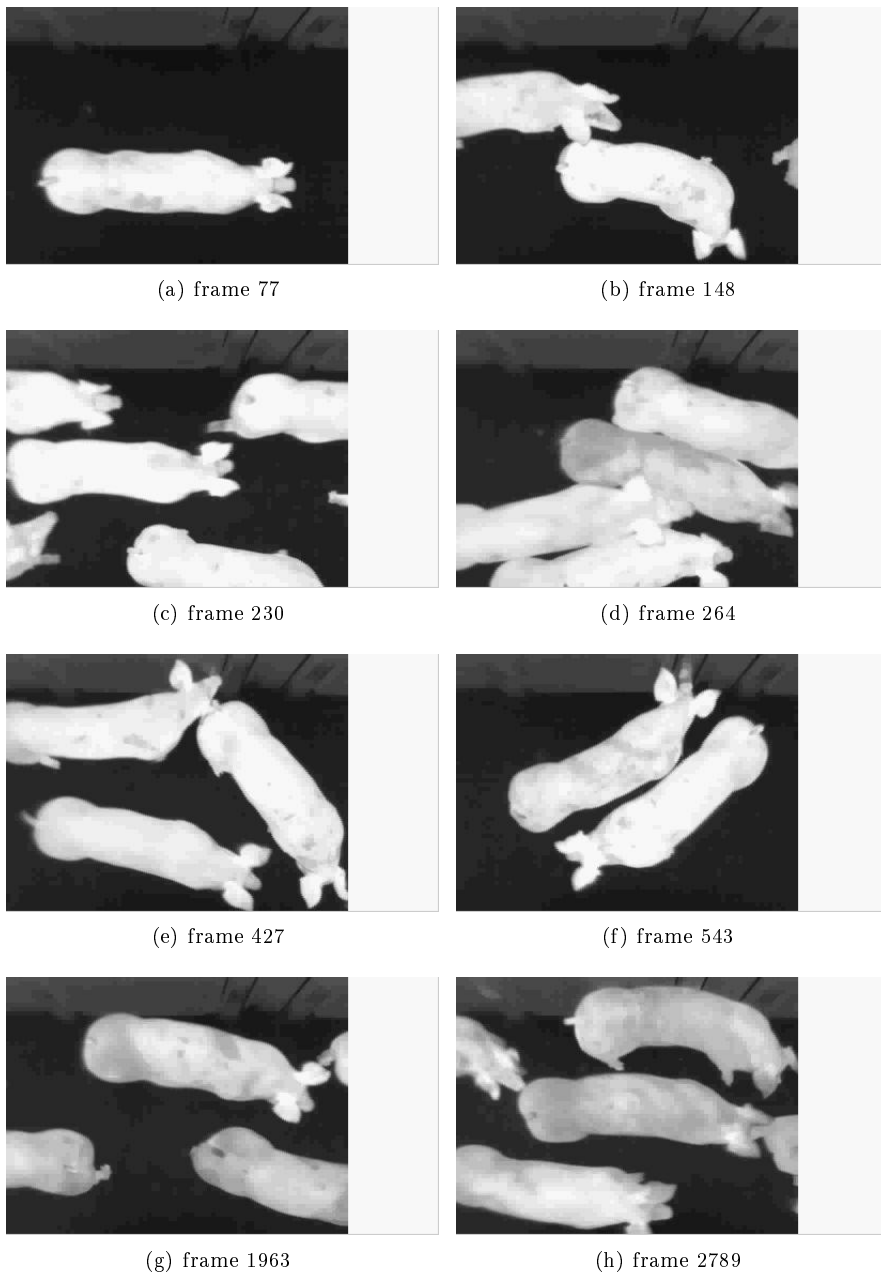
(c) frame 230

(d) frame 264

(e) frame 427

(f) frame 543

(g) frame 1963

(h) frame 2789

Figure 4.2: New received data.

(a) frame 2419

(b) frame 2451

(c) frame 2527
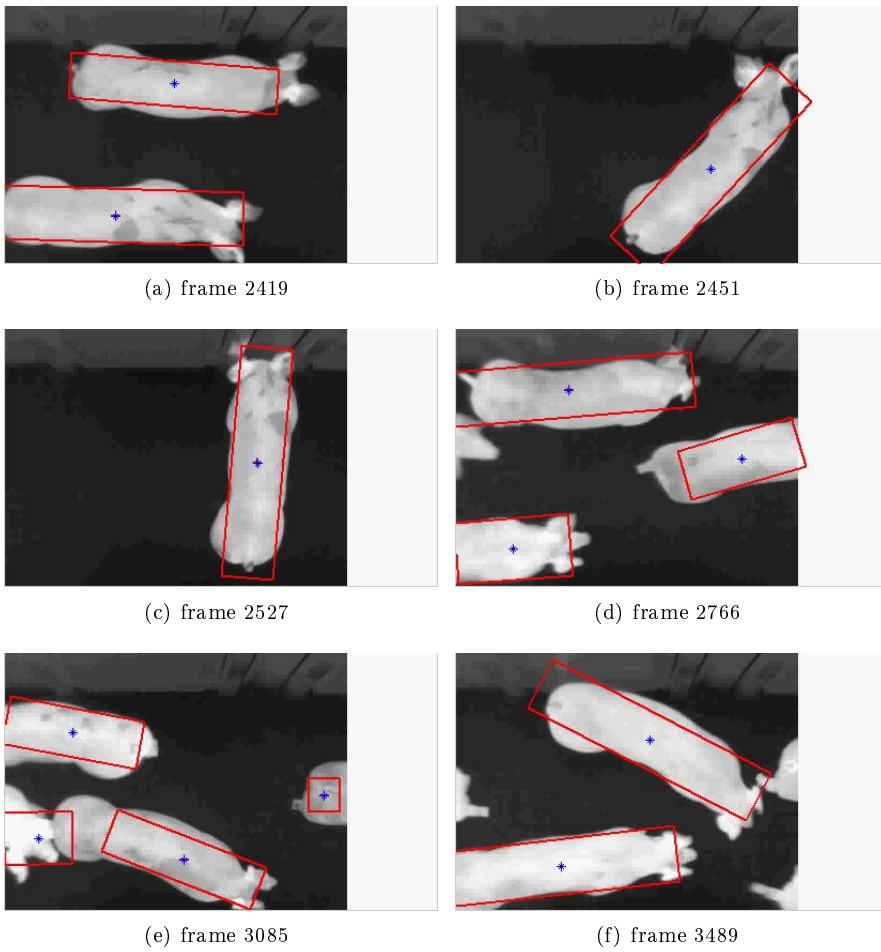

(d) frame 2766
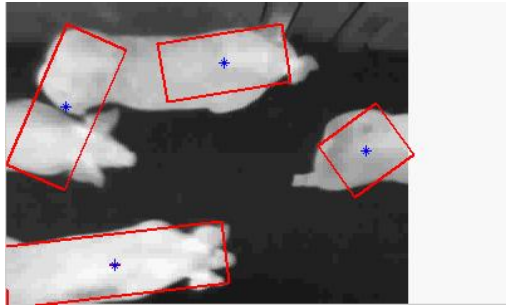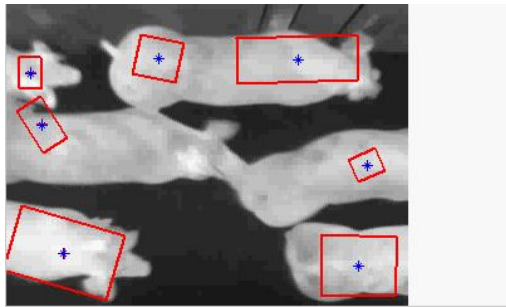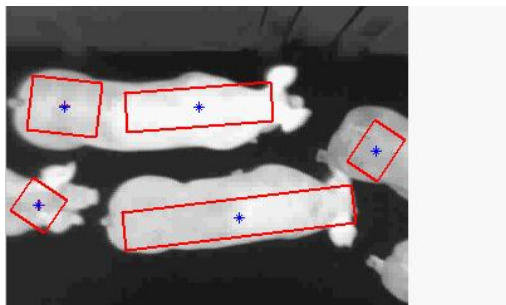

(e) frame 3085

(f) frame 3489

Figure 4.3: Output of the algorithm on the new received data.
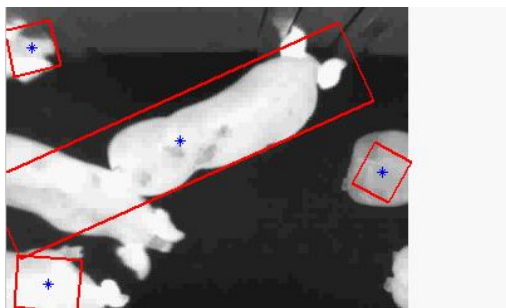
(a) frame 2769

(b) frame 2783

(c) frame 2948

(d) frame 3294

Figure 4.4: Some examples of errors on the output of the algorithm on the
new received data.

CHAPTER 5

Conclusions

5.1 Recommendations, possibilities and other considerations

According to the vets, the movement of the pigs can give enough information about the behaviour and the stress level of them. For this reason, locating the pigs on the image is very important. In this way, the tracking algorithm can provide a large amount of possibilities.

As I have been observing on the given video sequences, the implemented algorithm makes easier to automatise the detection of a pig that is impeding the free pass on the corridor to the other pigs. A pig that is turning around and pigs that are walking on the wrong direction can also easily be detected automatically.

Going beyond what the algorithm is, if the images had colour information, the extracted data could be larger. We know that a signal of abnormal behaviour on the slaughterhouse could be a pig biting another one. Of course, this behaviour leaves marks on the pigs skin, and then, these could be easily detected if they were not in a gray scale image.

Another strong point is the configuration of the camera location. My recommendation for the best tracking algorithm facilities is to locate a fix camera

some point on the roof of the corridor of the pigs, trying to have no perspective on the images and giving the maximum quality of them. Moreover, as I have said, if the images could include colour information, the possibilities for a successful tracking could be higher. Finally, the contrast is something important to differentiate the background to the foreground (pigs), but always having a compromise. If the contrast is too elevated, in crowded situations the task of isolating, detecting and tracking each pig becomes very hard.

To sum up, knowing what you are looking for makes the output information more reliable. Until now, the vets have not given a lot information about the pigs behaviour. This makes the task of giving a tool to find something you do not know difficult. Of course the clues that they gave us were more than enough to start an algorithm, but probably with more information the algorithm could turn to a more specialized method, and be more focused on finding the abnormal behaviour of the animals instead of just track them. That is something to take into consideration for the next steps of the project.

5.2   Conclusions

First of all I would like to put into consideration the importance of the video sequence characteristics to have a reliable data. As it has been shown in the last chapter, the results of the investigation are highly dependant on the choice of an appropriate video for the objective being tracked. On the second received video sequences, all the other problems that have been difficulting the tracking task have been easily solved with a more simple method than on the first case. Then, when the data are appropriate for the task that is being followed, it becomes much easier to solve. Therefore, as I have proposed on the last section 5.1, the choice of the correct input data is one of the most important steps before starting any tracking algorithm.

According to the observations of the final results, it can happen that some of them are found not very encouraging. The work behind all these results has been very hard, and it is important to know that all the work done with the algorithm has been very useful. Now it is just time to find the appropriate method to give a solution for the already found problems for the given images and apply it into the already done structure of the video sequences.

The last results obtained with the new received data are very encouraging. The proof that the work done with the tracking algorithm was good is the results obtained in less than a couple of weeks. Due to the incoming deadline of this project I had no more time to dedicate on the evolution of the algorithm. Because of that, I am sure that this results can be improved very much, arriving close to the ideal tracking.

In my case, let me stress that due to the deadline to hand in this project, it has been a hard task to have to stop looking for more results due to the deadline to hand in this project.

Poster for the Vision day

The Vision day is a one-day multi-track conference that covers a range of topics including computer graphics, machine vision, multivariate analysis, and image analysis for medical and food applications. I participated on the Vision Day on 30 of May doing the following poster about this project.

Analysis of well-being of pigs by means of video sequences

INTRODUCTION

The increasing demand of the alimentary industry for having an improvement of the quality in their products has become an opportunity to investigate and intercede in the meat industry.

Moreover, we have to bear on mind that the animal welfare is in focus, and handling of live animals before slaughter is one of the subjects for constant attention and discussion. The Danish Meat Research Institute (DMRI) has decided to combine this facts to start a research in this area.

The DMRI has large competences within animal welfare. They develop cost-effective solutions to improve animal welfare during transport, lairage and stunning, and they assist in implementing procedures and systems that contribute to proper animal welfare at the day of slaughter. Furthermore, they develop systems and methods to control and document the level of animal welfare towards customers and authorities.

In this case, the Danmarks Tekniske Universitet (DTU) collaborates with the Danish Meat Research Institute in order to implement an automated method which takes care of the well-being of pigs. The method will try to survey the well-being of pigs during the process from when they arrive at the slaughterhouse, when they are unloaded and finally they go to the CO^2 chamber. By means of video sequences, the algorithm tries to implement an animal tracking. Afterwards, the information about them is stored and used for extracting features about the welfare of pigs. The DMRI is providing the features that have to be the guide for the algorithm implementation.

ENGINEERING ISSUE: EXTRACT MAXIMUM INFORMATION ABOUT THE PIGS LOCALIZATION

TRACKING ALGORITHM

Tracking the DAISY features: the main task of the tracking algorithm. It is based on a matching of the features extracted between consecutive frames.

WINDOW UPDATE: once we have the matching position of the new pig we have to update the window that is "following" the aforementioned pig.

AUTOMATIC WINDOW INITIALIZATION: update, births and deaths of outdated windows.

SOLVING THE TRACKING:

DAISY FEATURES

The descriptors are very fast and efficient to compute. It depends on histograms of gradients like SIFT and GLOH but uses a Gaussian weighting and circularly symmetrical kernels. This allows a great speed and efficiency for dense computations. 200-length descriptors can be computed for every pixel in an 800x600 image in less than 5 seconds while the SIFT features take more than 250 seconds.

MATCHING THE FEATURES

The algorithm computes the DAISY descriptors in each pixel inside the window. Each descriptor becomes a vector of one hundred positions that gives the information of the related pixel.
For each obtained descriptor inside the window, the algorithm looks for the most similar descriptor on the next frame

The matching vectors start at the pixel of the current frame and end at the new found position on the next frame. They are plotted on the current frame.
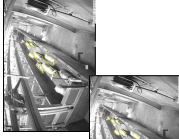
The similarity is calculated as the root mean square (RMS) of the descriptors of the two consecutive frames. The chosen pixel is the one with a minimum root mean square.

$$RMS = \sqrt{\sum_{i=1}^{100} (v_{b_i}(i) - v_{a_i}(i))^2}$$

WINDOW UPDATING

The window updating consists of displace the window to the new location of the pig on the next frame. The process is simple:

1. The angle of all the matching vectors is computed.
2. Then the mean of these values is extracted to know the mean direction of the window.
3. The corresponding matches that have an angle around the mean (variance of pi/15) are selected
4. Finally, the mean module of the selected matches is computed.

The mean module and the mean angle of the displacement of the window has been obtained. The window just has to be moved.
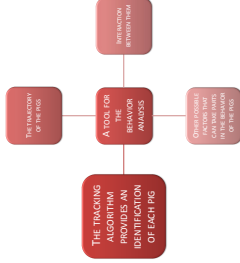
EXPECTED RESULTS

THE TRACKING ALGORITHM PROVIDES AN IDENTIFICATION OF EACH PIG

THE TRAJECTORY OF THE PIG

A TOOL FOR THE BEHAVIOR ANALYSIS

INTERACTION BETWEEN THEM

OTHER POSSIBLE FACTORS THAT CAN INDICATE THE BEHAVIOR OF THE PIGS

Author: Ricard Moya Mata (s114422)

Supervisors: Anders Lindbjerg Dahl and Bjarne Kjær Ersbøll

DTU

DMRI
Danish Meat Research Institute

Bibliography

[1] Slaughtering with maximum consideration for the animals. *Newsletter DMRI February*, 2010.

[2] Stefan Hinz. Fast and subpixel precise blob detection and attribution. In *Proceedings of ICIP, 2005*, pages 11–14, 2005.

[3] Tony Lindeberg. Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention. *International Journal of Computer Vision*, 11:283–318, 1993.

[4] Tony Lindeberg. Scale-space theory in computer vision, 1994.

[5] David G. Lowe. Object recognition from local scale-invariant features. *International Conference on Computer Vision*, pages 1150–1157, 1999.

[6] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004.

[7] Engin Tola, Vincent Lepetit, and Pascal Fua. Daisy: An efficient dense descriptor applied to wide baseline stereo. *IEEE TRANS. PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 32(5), 2010.