Development of a Teat Sensing System for Automated Milking

By

Aymen Ben Azouz

This thesis is being presented in fulfilment of the requirements for the qualification of Masters of Engineering

Dublin City University

Supervisors: Dr Harry Esmonde
Dr. Brian Corcoran

School of Mechanical and Manufacturing Engineering

Teagasc Advisor: Dr. Eddie O'Callaghan

Dairy Production Research Centre, Teagasc , Fermoy, Co. Cork

August 2009

Volume 1 of 1

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Masters of Engineering is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Aymen Ben Azouz

Signed:
ID No: 57101213
Date:

“O, my Lord! Advance me in knowledge.”

(The Holy Quran 20: 114)

Acknowledgements

This research would not have been possible without the continuous help and support of many people. I wish to express my heartfelt gratitude to everyone who have enabled and supported the work behind this thesis. I would like to acknowledge and thank my supervisors, Dr Brian Corcoran for his advice, constant support and encouragement, and Dr Harry Esmonde for his guidance, advice and ability to resolve issues I sought his help on.

I would like to thank Teagasc, The Agriculture and Food Development Authority in Ireland and specially Dr Eddie O'Callaghan for providing the funding and support to pursue this research work under the Walsh fellowship scheme.

I would like also to thank the staff of the Mechanical and Manufacturing Engineering School who provided the technical assistance and advice to achieve this work.

On a personal note, I would like to express my sincere gratitude to my colleagues and friends in DCU for their ideas, help and encouragement during the long hours spent on campus.

To my parents and family, for unquestioningly supporting and always encouraging me to be the best I can be. The self belief that they have instilled in me and the unconditional support they offered has provided the strong platform on which I build my life.

Finally, I would like to thank all people that have assisted, supported and inspired me, both directly and indirectly throughout my academic endeavours.

TABLE OF CONTENTS

ABSTRACT

Robotic application of milking cups to the udder of a cow in a rotary high capacity group milking system is a major challenge in automated milking. Application time and reliability are the main constraints. Manual application by an operator of a rotary carousel is of the order of 10 seconds and 100% reliable. In existing non-rotary milking machines, the cups are applied to each teat individually and the process can take up to two minutes. In order to achieve a more rapid simultaneous application of the four cups, the three dimensional locations of the four teats must be known in real time. In this thesis, a multimodal vision system combining optical stereovision and thermal imaging is developed. The overall system is evaluated from the point of view of accuracy and robustness. Laboratory tests have shown that stereovision can rapidly locate teat three dimensional position coordinates, however robust identification of the teats is required. It is shown that this may be achieved using thermal imaging to isolate teats from background objects due to their elevated temperature profile. Further development is necessary to overcome specific situations such as overlapping teats.

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1- Research motivation:

Dairy farming is one of the largest sectors in Irish agriculture. Dairy exports account for 75% of total agriculture production (Agriculture in Ireland, 2002). Due to economic and social pressures, automation of the milking process is becoming both desirable and necessary. Tasks such as cluster removal, teat spraying, milk monitoring, milk diversion, cow identification and health monitoring have already been successfully automated. Consequently the manual attachment of milking cups to each animal is the most time consuming element of the process for the farmer during milking.

Nearly all the dairy farming in Ireland is pasture based. Existing Automatic Milking Systems (AMS) are unsuitable for use with grazing because the higher milking frequency is difficult to implement. The farmers would need to install more than one AMS; this significantly increases the capital outlay involved for a farm moving to automatic milking. Also, the dependence on concentrate food offsets the savings incurred with grazing.

A modern rotary carousel, such as the one pictured in Figure 1.1, is suitable for milking large herd sizes. With a 60 stall rotary it is possible to milk up to 300 cows per hour (O'Callaghan et al., 2001).

Figure 1.1: A Rotary Carousel (DeLaval Rotary, 2008)

Automating the attachment of the milking cups on a rotary parlour has numerous advantages over the traditional automated systems currently available, and over a non-automated rotary parlour: Teagasc and DCU are investigating an AMS for application to a rotary carousel and have subdivided the research into two parts, that is teat sensing and

cup application. In terms of sensing, rapid, precise and reliable detection of teat position is essential so that the robotic cup applicator can function in real time.

In this work, a multimodal, thermal-stereovision, teat sensing system is proposed. This system augments traditional stereovision detection systems with reliable thermal imaging technology. The proposed system overcomes the limits shown by different systems tested for the purpose of teat sensing within an AMS and is able to handle:

1. Interference of other objects in the scene, even when it is a teat shaped object
2. Variability of configuration of the teats angulations

1.2- Automation in Irish Dairy Farming

Dairy farmers need to reduce milk production cost in order to compete with reduction in milk prices. This can be done by increasing levels of milk production. However, to increase the production, dairy farmers need to enlarge the size of their herds. Larger herds mean increased labour and as Figure 1.2 shows, the milking process is responsible for one third of the total net labour input associated with dairy farming.

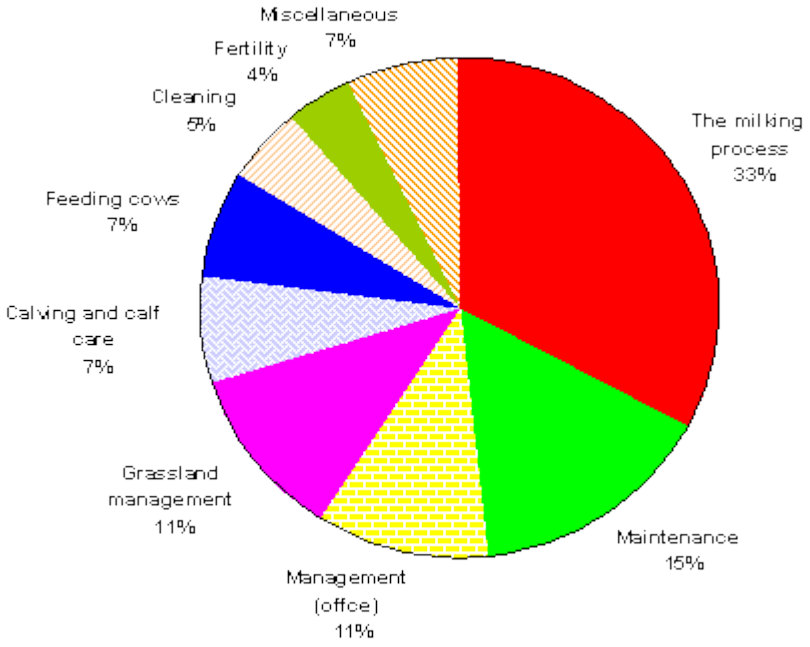


Figure 1.2: Breakdown of total net labour input associated with dairying over the 12 month period (O'Brien et al., 2002).

Therefore, the milking process would be the area to focus on in order to make labour-cost savings. In addition, small dairy farmers, who cannot make cost reductions, have ceased production of milk and their quotas have been taken up by the bigger dairy farms. This creates a demand for additional labour on the larger farms, however, dairy farming is

considered as a hard-labour sector and it has become very hard for dairy farmers to find operatives willing to work seven days a week, 52 weeks a year for a relatively low wage. The development of the industry in Ireland in the last couple of decades has pushed people from dairy farming to other industrial sectors where they can have more social working hours and better wages. Experts agree that the most obvious way to fill the void between the workload and the available labour on a dairy farm is to automate the milking process (O'Callaghan et al., 2001). As well as reduction in milk production cost, automating the milking process creates more free time for the dairy farmer. According to a survey, among 107 farmers who have recently invested in an AMS in mainland Europe, two thirds of the interviewed AMS farmers state social reasons for investing in an AMS (Mathijs, 2004). The social reasons they enumerated are reduction of labour intensity, dedication of more time to other activities, avoidance of health problems related to milking activity, increased challenge for their daily routine, improved social life and animal welfare (Wade et al., 2004; Jensen, 2004). AMS also help the farmer in others tasks such as udder cleaning, improved cow and udder health and improved management information such as better traceability (Mottram, 1997; Mathijs, 2004).

1.3- Existing Milking Techniques:

There are two types of milking systems:

- Voluntary milking system
- Parlour milking system

1.3.1- Voluntary milking system:

Voluntary milking systems are gaining widespread acceptance, and by 2004, it is estimated to be in use on more than 2200 farms in over 20 countries worldwide. More than 80% are located in north-western Europe (K. Koning, 2004). In voluntary milking systems (VMS) all tasks related to milking such as teat position location, teat cup application, cluster removal, teat spraying, milk monitoring, and milk diversion are automated. VMS are designed for milking permanently housed cows. The VMS are in operation 24 hours a day, 365 days a year. The cows are milked on demand so that it is up to the cows to decide when to go to the VMS to be milked. The cows are attracted to the VMS by providing concentrate feed in the milking stall. Studies have shown that on average, a cow is milked around 2.6 times a day. The incorporation of a VMS to a dairy

farm is costly, the system itself costing around €150,000 per two stall unit. Further costs are incurred through barn modifications, the culling of cows that are incompatible with the AMS, and the regular maintenance of the system (Review Document, 2006). The introduction of VMS to a farm generally results in a reduced production of milk (due to the restricted herd sizes). This is explained in the results of the studies described in (Mathijs, 2004) and (Wade et al., 2004); Dairy farmers decide to move to VMS to improve their lifestyle rather than to save money.

DeLaval VMS (voluntary milking system) is the current world market leader (DeLaval, 2008).

Figure 1.3: The DeLaval VMS (DeLaval VMS, 2008)

Other systems include the Lely Astronaut A3 robotic milking system and Lely Astronaut A2 robotic milking system from Lely Group (Lely, 2008) and the MERLIN AMS from Fullwood Limited (Fullwood, 2007).

1.3.2- Parlour milking system:

There are a few types of parlour systems on the market. However the most developed and widespread are the herringbone parlours (Figure 1.4) and the rotary parlours (Figure 1.5).

Figure 1.4: DeLaval Endurance herringbone Heavy Duty (DeLaval herringbone, 2007)

Figure 1.5: Rotary Milking Parlours from Dispatch & Garlick Ltd, New Zealand (Dispatch&Garlick, 2009)

In the last few years some degree of automation has been introduced into parlour based milking. In countries where cows are pasture- based like USA, Australia, New Zealand and recently Ireland and UK, dairy farmers are modernising the milking systems. In New Zealand for example, over the past 30 years, dairy farmers have refined batch-milking to

extreme levels of efficiency with large herringbone and rotary dairy systems currently accepted as best practice (Woolford et al., 2004). The national average herd size is continuously increasing (Figure 1.6) and there has been a steady shift in the type of parlour installed. A rapid uptake of herringbones occurred between 1973 and 1984 and the herringbone continues to be the most popular type of milking system in New Zealand. However, herd size has a major influence on choice of parlour type with larger herds having a relatively higher percentage of rotaries than smaller herds (Table 1.1).

Rotary milking systems vary in size from 30-100 stalls, the median being approximately 50 stalls. These milk harvesting systems achieve extremely high cow throughput, with minimal operator time, often with as little as 5-6 seconds for cluster attachment per cow. 'This high intensity batch-milking scenario is now the norm for New Zealand dairy farmers and they are arguably the best in the world at rapid milking of large numbers of cows with minimal labour while at the same time achieving excellent levels of milk quality (Woolford et al., 2004).

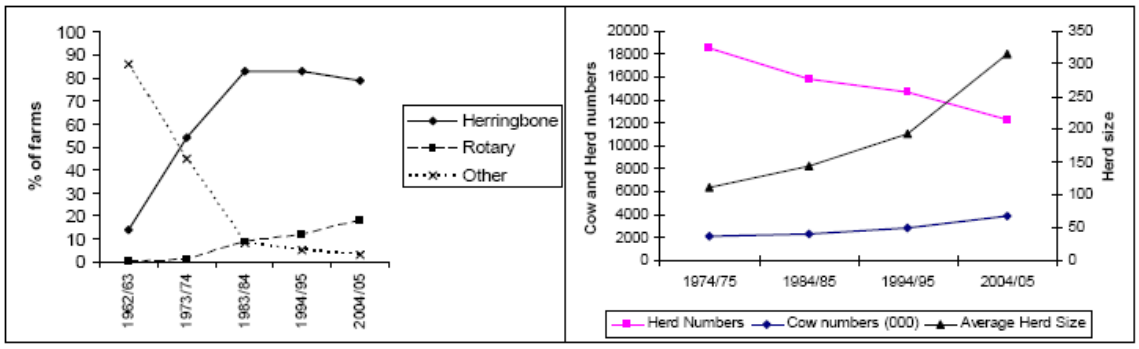


Figure 1.6: Percentage distribution of Milking Parlour Types in New Zealand (1973 to 2005) and Total cow and herd numbers and herd size in New Zealand, 1974 to 2005 (Jago, Ohnstad and Reinemann, 2007)

Herd Size(Cows)	**<100**	**100-199**	**200-299**	**300-39**	**>400**
Herringbone	67	88	82	6	53
Rotary	3	7	17	31	40
Walk through	25	3	1	0	0
Other	5	2	0	3	7

Table 1.1: Distribution (%) of parlour type by herd size in New Zealand (NZDB, 1984)

The conclusion that emerges from these studies is that the rotary milking parlour is suitable for milking large herd sizes. Figure 1.8 explains the rotary milking process; cows are kept in the enclosure adjacent to the rotary carousel until rotation of the carousel begins. The gate is then opened to allowing cows to pass in a single file to the carousel entrance during which sensing equipment identifies each cow by reading information stored electronically on a tag attached on the body of the animal (usually the ear). If cleaning of teats and udder is required it can be carried out automatically as the cow approaches the carousel or an additional labourer can do it manually. As the cow enters the empty stall in the rotary, the stall is automatically closed off at the rear to prevent the cow from backing out. The carousel rotates so that the next empty stall becomes available to the succeeding cow and the process repeats. The operator stands downstream of the rotary entrance so that the rotation of the carousel brings the cows towards him (Figure 1.7). Standing below the level of the carousel floor, the operator has easy access to the teat cup liners and to the udder of the animal. An audio alert system can be used to indicate to the farmer the identity of the next cow that is approaching him. Information regarding mastitis or antibiotic treatments can be issued and the farmer can leave certain quarters un-milked accordingly. The speed of the carousel is controlled so that the cow is completely milked as the carousel completes one full revolution at which point the clusters are removed automatically and the rear of the stall opens allowing the cow to back off the carousel (Review Document, 2006).

Figure 1.7: View of the manual teat application with rotary carousel

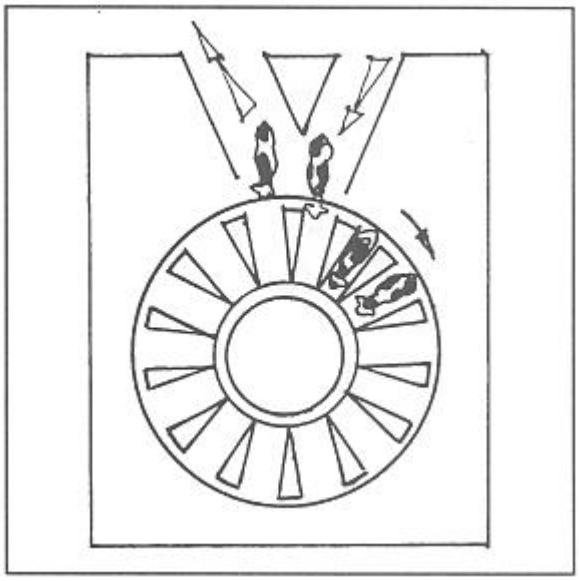


Figure 1.8: Schematic of rotary milking process (KwaZulu, 1999)

1.3.3- Rotary parlour and Irish dairy farming:

In Ireland, nearly all dairy farming is pasture based. VMSs are not suitable for use with grazing due to difficulties arising when trying to achieve frequent milking. Grazing, being an abundant supply of food, causes the lure of the concentrate in the milking stall to lose its importance. More than one VMS unit is required in order to achieve optimum milk yield and the extra VMS spread out over the grazing area reduces the distance that cows are required to walk to be milked, making the process more appealing. The drawback of this strategy is that the capital outlay involved for a farm moving to automatic milking is significantly increased. Apart from the cost of the additional VMS units, other factors must be taken into consideration such as extra milk lines, power lines and cooling. Maintenance and monitoring of cows becomes more difficult due to the distance between the units and becomes more time consuming for the farmer. In addition, the savings incurred with grazing are offset due to the dependence on concentrate food. Allowing the cows the freedom of a natural grazing habitat will inevitably cause the cows to settle into eating and milking habits influenced by the environment. This will cause peaks when the VMS is inconstant operation and troughs when it is idle. The main justification for the huge cost of installing a VMS is the fact that it is in constant use 24 hours a day, 365 days a year, which is never the case with grazing. Therefore, the VMS units create a source of inefficiency by not being utilised to their maximum capacity on the dairy farm.

By contrast, the rotary topology is suited to the conditions of Irish dairy milking and once automation of the attachment of the milking cups is introduced, the system would offer a

time efficient milking solution. The use of a fully automated rotary parlour will have significant advantages over existing systems such as

- Greater Capacity: The throughput of the system is not restricted by the milking time of the animal as with herringbone systems. Increasing the number of stalls in the carousel and reducing the teat cup attachment time could increase the capacity.
- More Cost Effective: A single robot can be used to service more cows. There is a better use of assets than with the VMS technology.
- Reduced Labour: The farmer is only required to herd the animals to the AMS Rotary. Presence is only required when a problem occurs. This frees up a significant quantity of time for a farmer who would usually have to be present for the entire milking process using a conventional rotary system (manual attachment of teat cups).
- No Down Time: In case of an AMS malfunction and required maintenance the herd will not go un-milked. Should the system be non-operational the rotary system gives the farmer the possibility to milk the herd by manually attaching the teat cups to the cows.
- Conventional rotary parlours are already established in the Irish dairy industry. These have proven successful and farmers are more willing to consider modification based on a tried and trusted system.
- The approach to dairy farming does not change as drastically as it does when a VMS is introduced. The farmers can continue to use the food sources that they are currently using, which mean that grazing is not affected.

1.4- Teat sensing:

One of the main challenges in the AMS is determination of teat position. The teat positions differ not only between cows, but also in between consecutive milkings of the same cow. The teat location can be influenced by differences in udder shape and condition, inner pressure of the udder, but also by the length of standing or laying time (Bergerot, Baylou and Ordolff, 1989). If the animal moves during the measurement or application phase, real time updating of position is required. This paragraph reviews, first, existing teat sensing systems within commercial VMS and secondly, some research work attempted in the past to design such teat sensing systems.

A teat sensing system should satisfy the following requirements (Artmann, 1997):

1. The determination of the space coordinates must be exact and fast enough for the application to be successful.
2. Invasive treatment of the animal (implants, etc) should not be required.
3. The animal should not be touched during the location process in order to avoid provoking uncontrolled movements.
4. The sensor system must be appropriate for the particular environmental conditions at the implementation site, and be robust and reliable

1.4.1- Technologies used in existing VMS:

While the parlour milking systems are not equipped with milking cups attachment systems, the VMS are fully automated system and incorporate a teat sensing and cup attachment system (Internal Report, 2006; Review Document, 2006).

All sensing systems used in the commercial VMS are laser/vision based systems. A single camera is used for the detection of the laser stripe location that is incident on a teat. In order to triangulate the position of the teat, the location of the stripe in the camera image is coupled with the relative distances between the laser and the camera and the angles of incidence.

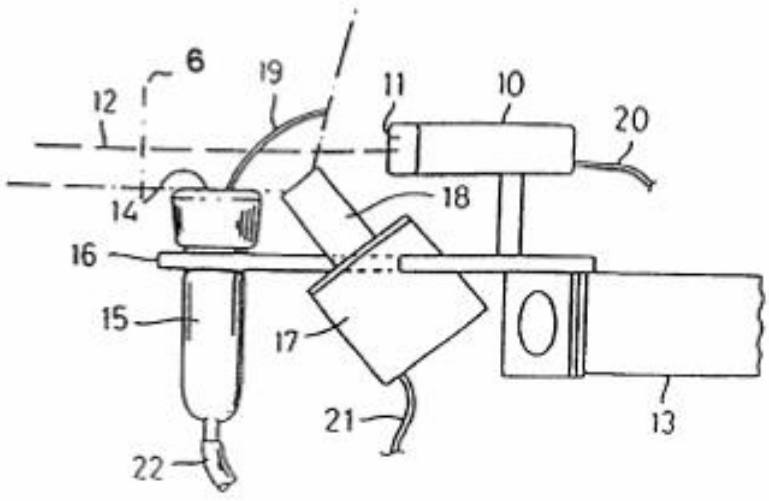


Figure 1.9: Image from patent WO98/47348: Single Camera and Laser stripe generator

Note: Figures 1.9 and 1.10 appeared in the international patent WO90/47348 (Andersson and Nilsson, 1998).

Figure 1.9 shows a camera (no. 17) and a laser horizontal stripe generator (no. 10) attached to the end-effector of a manipulator arm (no. 13). The end-effector is capable of gripping the milking cup (no. 15) by means of a carrier (no. 16). The end effector is

moved beneath the cow in order to scan the laser across the udder. When the horizontal laser stripe is incident on a teat a line segment will be seen in the image of the camera. Figure 1.10 shows a series of line segments in the image corresponding to the teat; this is obtained by scanning the end-effector in a vertical plane. The ends of the segments define the outer edges of the teats. When the teat is identified, it is possible to determine its position by using laser triangulation techniques.

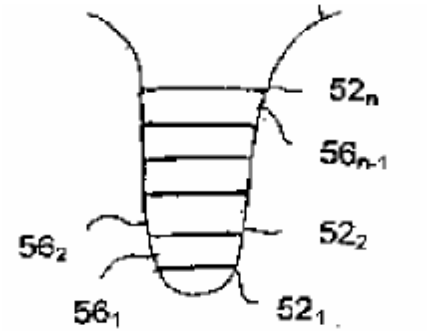


Figure 1.10: Image from patent WO98/47348: Series of illumination stripes on a single teat

Such a system can provide reliable and accurate information regarding the position of a teat but it can only detect a single teat at a time. Once a teat has been successfully located and the milking cup is applied the system moves on to the next teat. This procedure is repeated until all four milking cups are attached. This method is not a suitable means of teat detection for herd-wise milking due to the time overhead associated with the scanning motion. For the current approach, the positions of all four teats are required simultaneously.

1.4.2- Teat sensing research:

First attempts to locate teat position date back to the period of introduction of the first components for controlled milking (Gabler, 1971; Notsuki and Ueno, 1977) in the seventies. However, successful solutions could not be achieved prior to the availability of powerful electronic sensing and controlling devices at moderate prices. Fundamental work on automatic milking was done at IMAG-DLO in the Netherlands, at CEMAGREF, France, at AFRC Institutes at Compton and Silsoe in Great Britain, at the Federal Dairy Research Centre at Kiel (Germany) and at the Agricultural Research Centre (FAL) at Braunschweig-Volkenrode (Germany) (Ordolff, 2001). In one system, a circular device divided in sectors equipped with conductivity sensors (Akermann, 1979), was used to

detect the position of a teat see Figure 1.11. Later, mechanical contact sensors were developed for the same purpose (Artmann and Schillingmann, 1990). However, these techniques were quickly dropped as mechanical sensors can hurt the animal in addition to the unwanted movement of the cows provoked by these sensors (Artmann, 1997). During the eighties, successful attempts using image processing with laser scanning techniques (Montalescot, 1987) and ultrasonic (US) range finding devices and/or optical sensors were realised (Torsius, 1987; Scheidemann, 1990). A number of processes are available for US distance measurement (Artmann and Schillingmann, 1989; Artmann et al., 1990); the exact teat location is determined by two sensors from the sensor array, which is arranged so that the US search region from the two sensors cross each other in the teat region so that triangulation is used to calculate the teat position. Laboratory tests showed that with this arrangement in the intersecting US waves, it is possible to measure to an accuracy of about 1 mm. This method was used in the Duvelsdorf AMS (Duck, 1992). Greater success was achieved with these systems in conjunction with a data base of relative teat configurations. Historical knowledge of relative teat position helped to reduce the amount of time needed to identify all four teats consecutively. However, as explained in the previous paragraph, current VMS use laser scanning techniques. While the detection with laser scanning techniques is faster than US, neither are fast enough to integrate into a fully automated milked system for teat detection on a rotary parlour.
Thermal characteristics of the udder region were also explored (Ordolff, 1984), however, at the time thermal technology was very expensive and very slow.

Stereovision techniques have also been explored. Stereoscopic processes require pictures taken from two positions. This can be accomplished with one camera when the camera is moved a defined distance parallel to the first, and the second picture is taken from the new position. It also can be accomplished with two cameras at a defined distance from each other. With moving objects (cow) the second method is more advantageous. One of the main problems with the use of stereoscopic processes in AMS is a quick and reliable solution for the correspondence problem. This means that the corresponding point in the same scene must be found in both pictures [33]. Xu and Aneshansley (1991) implemented and demonstrated such a system by attaching teat cups to the teats on an artificial udder (Artmann, 1997; Ordolff, 2001). More recently, and within the Teagasc/DCU collaboration, a Teat Tracking Vision system, using stereovision has been assessed (Duffy, 2006). A summary of the different techniques used in teat sensing research is shown in Figure 1.12.

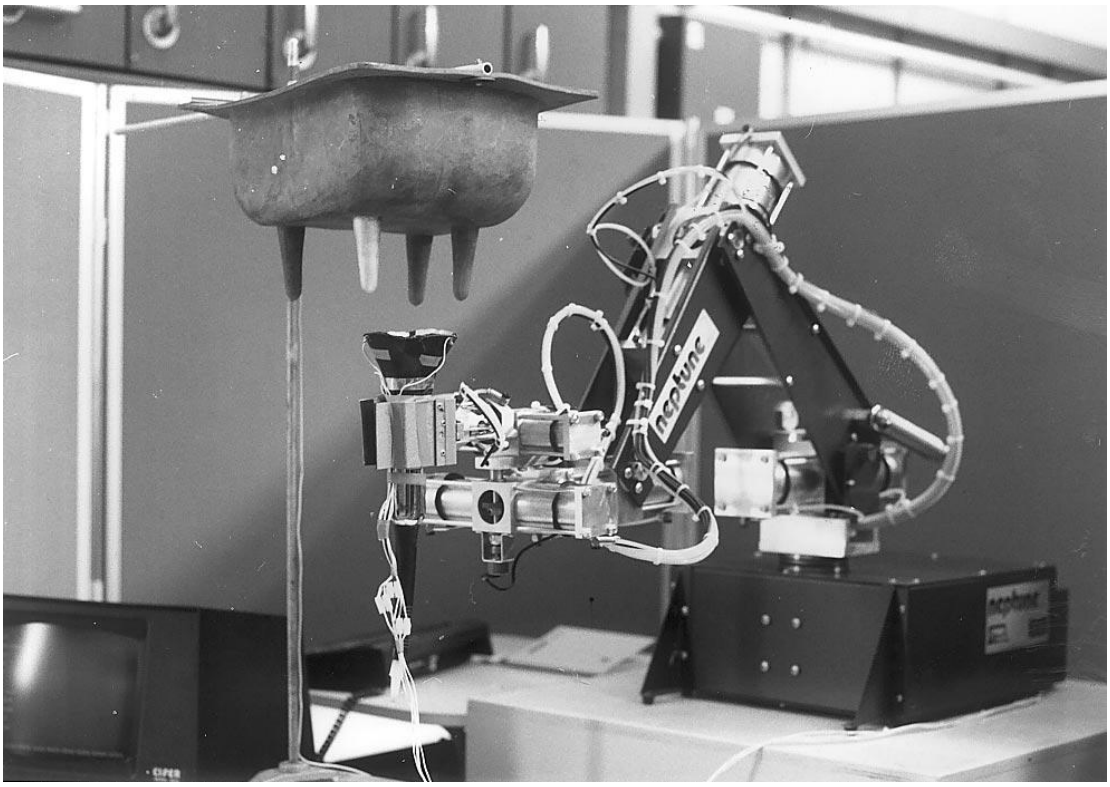


Figure 1.11: Experimental device for teat localization by a galvanic sensor (Ordolff, 2001)

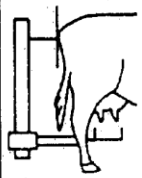
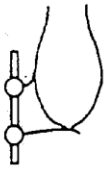
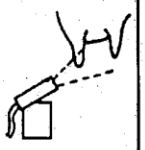
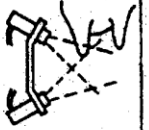
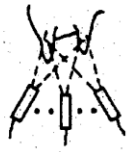
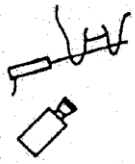
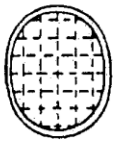
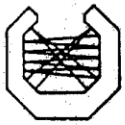
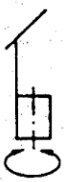
sensors	Teat detection						
	Contact with cow	2 tactile sensors	1 US sensor for scanning	2 US sensors for triangulation	US sensor array	Image processing mobile laser	Image processing fixed laser
global							
local US=Ultra sonic	large orifice of teat cup mouthpiece	light beam matrix	light beam matrix	rotating US sensor	2 US sensors for triangulation	light beam matrix	—
							—
Company	Gascoigne-Melotte	AFRC	Düvelsdorf	Vicon	FALI	CEMAGREF	FAL II

Figure 1.12: Sensors for teat detection in automatic milking systems (Artmann and Schillingmann, 1990).

1.4.3- Proposed solution:

Manual application of the cups by the operator of a rotary carousel is of the order of 10 seconds and of course 100% reliable. As a design goal the automated attachment should not take any longer than 20 seconds to maintain high productivity. Considering that the cup attachment task in itself depends on the speed of the robot the teat location task must be rapid. Although cow movement during the attachment process is limited, there is a need for real-time detection so that the system refreshes the positions of the teats every one to two seconds. Therefore, considering this condition as well as the requirements of reliability and accuracy, the teat sensing system must have:

1. Time detection of one second or two seconds maximum
2. High reliability of 100% or near 100%
3. An accuracy of 5mm (detection error less than 5 mm)

Despite their robustness and high accuracy, laser scanning techniques cannot be considered as teat sensing technique for an attachment system in conjunction with a rotary parlour. The laser scanning system used with commercial VMS takes up to two minutes to complete the teat cup application stage. Such systems are too slow and would affect the throughput of the rotary milking process.

Ultrasonic (US) technology has been also used in previous attempts. However, the robustness of these systems was low. In addition, the use of ultrasonic waves within a confined area, such as the area underneath a cow, would generate difficulties in processing US signals due to multiple reflections.

Stereovision is a rapid means of detection and location of objects in 3D space. With modern high speed processors, identification and location of the teats of cows with two cameras using stereovision can be realised within a few tenths of a second. Stereovision have been used in research as a means of teat sensing since the eighties, however, at that time, digital imaging was not developed to its current level.

During previous work on this project (Duffy, 2006), the performance of a commercial stereovision system was assessed in two ways, namely.

1. The success in identifying teats in the images.
2. The accuracy of the system in determining teat position.

1.4.3.1 - Teat tracking vision system: Identification assessment:

The results showed that a successful identification occurs when the system correctly identifies corresponding image points in the left and right images of the stereo pair belonging to the same teat target. Teat identification works well under ideal conditions where the scene contains four unobscured, well-illuminated teats, in a standard udder formation, located within the target region of the vision system and against a neutral background.

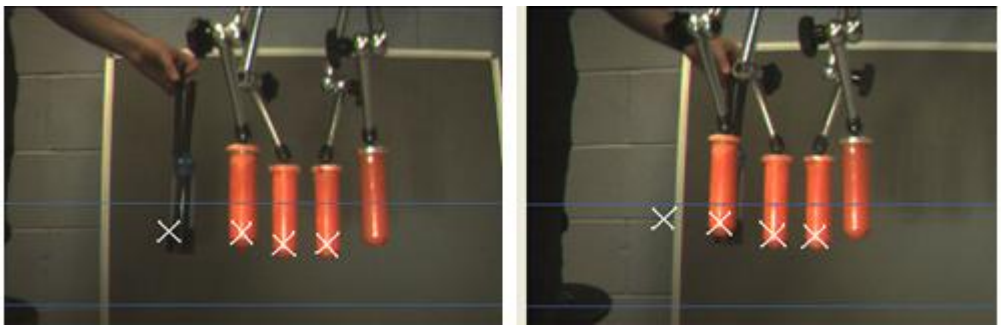


Figure 1.13: Teat incorrectly identified

With teats partially obscured or occluded the identification process breaks down. It was found necessary to have the bottom of the teat clearly visible in both camera views for proper identification. Presence of other objects in the cameras' fields of view severely reduces the system ability to extract the teats from the image. In Figure 1.13, the "spurious" teat in the background of the image is picked up instead of the far right teat.

Of course further refinements could be made to improve the identification process in the example shown in figure 1.13 above but teat misidentification was found to occur for several reasons such as: when movement occurred (particularly in the background), when lighting parameters changed and also when no apparent change had occurred .

1.4.3.2- Teat tracking vision system: Accuracy of teat location:

The accuracy of the system was tested by taking measurements of a single teat moved to different locations within a 300mm by 275mm horizontal plane. These measurements were compared to known accurate measurements for the teat locations, see Figure 1.14. The 3D surface plot clearly illustrates the error of the systems measurements throughout the test plane. The region of best accuracy is achieved in the central area, as the teat location moves away from this area the error increases. The average error for the entire test region is 11.79mm. The largest error is in excess of 15mm and the smallest error is less than 1mm. At the outset of this project it was proposed that a teat location

measurement with a certainty of ±5mm or less would be sufficient in autonomously attaching a milking cup to the teat. To attain an average error within this requirement the test region must be confined to an area of 150mm wide by 80mm deep. Only the minimum expected teat formation dimensions fit inside this area. It is therefore not possible for the system to simultaneously provide four accurate location measurements for teats in the average standard udder formation. Either the front teats or the rear teats will be outside of the region of an accuracy of ±5mm (Duffy, 2006).

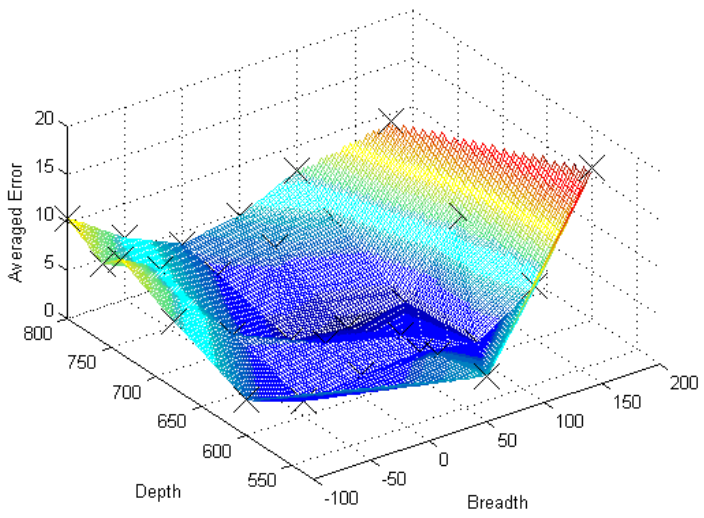


Figure 1.14: Identified teat position error

If the system was modified so that the ±5mm accuracy area was increased to encompass the maximum expected teat formation dimensions, it would still not be sufficiently large since the animal can move to certain extent within the stall.

1.4.3.3- Integration of stereovision with thermal technology:

The udder and teats have a greater external temperature than other regions on the surface of the cow or typical backgrounds. Thermal imaging is proposed as a possible means of reliable teat identification.

Thermal imaging has developed in last decades. Research with thermal imaging have been mostly carried out in the military field (Newman and Hartline, 1982) and for medical applications, such as a non-invasive diagnostic tool (Jones and Plassmann, 2002) for maladies such as arterial disease and for therapy monitoring (Burnay, Williams and Jones, 1988). Thermal infrared imaging has been shown to be a very useful modality in many areas of computer vision research. As thermal radiation comes primarily from emitted radiation, it is very robust to adverse lighting conditions that plague many visible

spectrum vision systems. It is also very useful for distinguishing people and vehicles from complex backgrounds. Some drawbacks of thermal imaging that may limit its use are:

- high levels of image noise
- the halo effect in ferroelectric cameras
- inability to distinguish objects of similar temperature, such as people
- Low pixel density in thermal imaging hardware by comparison with optical systems

Consequently a combination of thermal and optical imagery will be used for teat identification. The thermal image will be used as a robust means of teat identification, extracting the teats from spurious background objects. This information will be passed on to the optical system thereby significantly reducing the search region used to find point correspondences. The high pixel density in the optical images will then permit accurate determination of teat end position. (Ben Azouz et al., 2008).

CHAPTER 2

Theory

This chapter provides a brief description of the theory behind some of the techniques used throughout this project. The first section deals with multi view geometry (which underlies the theory of the projection of points through the camera centre), the techniques of stereovision and the calibration method used in this work. The second section reviews some of the image processing techniques used to detect the ends of the teats.

2.1- Multi view geometry and computer vision:

The first two paragraphs of this section explain the geometry of single and double view vision systems. In the third paragraph, a review of camera calibration will be presented and explained.

2.1.1- Single view geometry:

2.1.1.1- Pinhole camera model:

A camera projects points from 3D world coordinates onto a 2D image (Hartley and Zisserman, 2003b). The pinhole camera model is the most basic model used to describe this transformation in computer vision applications (Sonka, Hlavac and Boyle, 1999). The pinhole camera model assumes a camera with an infinite, thin lens, so that all light is focussed through a single point in space, the camera centre. Such lenses map each ray of light to a unique point on the image plane (Hartley and Zisserman, 2003a). Figure 2.1 describes the geometry of the pinhole camera.

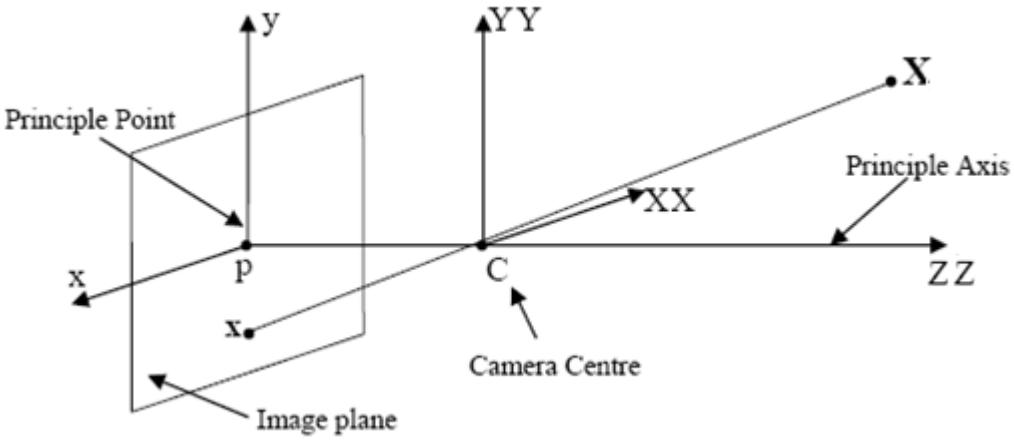


Figure 2.1: Pinhole camera model (Duffy, 2006)

In figure 2.1, C is the centre of the camera, also known as the focal point. The principle axis, which is the perpendicular line to the image plane passing through the point C, intersects with the image plane at p, known as the principal point of the camera. The distance pC represents the focal length of the camera, f.

The projection of point X is represented by the optical ray (ray of light) Xx. The ray passes through the centre of the camera C and projects the point X onto the image plane at x.

From the previous presentation of the pinhole camera model, four coordinate frames can be defined (Sonka, Hlavac and Boyle, 1999):

1. The camera Euclidean coordinate frame: has as centre the camera centre C and the orthogonal axis system XX,YY,ZZ
2. The world Euclidean coordinate frame: for this example, the world Euclidean coordinate frame is the same as the camera Euclidean coordinate frame. Usually, this frame can be defined anywhere in space. The transformation between these two frames consists of a rotation and a translation.
3. The image Euclidean coordinate frame: this frame is related to the image plane, it has as origin the point p and axes x and y. The z axis can be defined as the perpendicular to the image plane emanating from point p in the opposite direction to Z. However, this coordinate is not assigned to the points in the images which are regarded as 2 dimensional.
4. The affine image coordinate frame: it is a frame related to the image plane, with p as origin and axes related to the characteristics of the pixels of the camera. This extra frame arises because, in general, pixels can be non square.

2.1.1.2- Camera parameters:

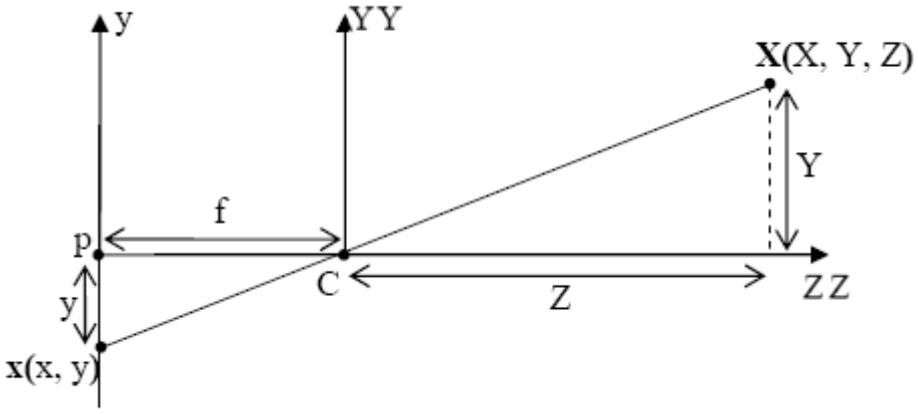


Figure 2.2: Pinhole projection in the YYZZ plane (Duffy, 2006)

From Figure 2.2, considering the plane YYZZ:

$$\frac{y}{f} = \frac{Y}{Z} \tag{2.1}$$

$$\Longrightarrow y = fY/Z \tag{2.2}$$

Similarly:

$$x = fX/Z \tag{2.3}$$

Therefore the perspective projection that applies for the pinhole camera model from the 3D scene to the 2D image plane can be defined as:

$$(X, Y, Z)^T \rightarrow (\frac{fX}{Z}, \frac{fY}{Z})^T \tag{2.4}$$

Using homogenous coordinates, this projection can be expressed as:

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} fX \\ fY \\ Z \end{pmatrix} = \begin{bmatrix} f & & & 0 \\ & f & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{2.5}$$

The matrix $P = \begin{bmatrix} f & & & 0 \\ & f & & 0 \\ & & 1 & 0 \end{bmatrix}$ is the camera projection matrix, known also as the camera matrix (Hartley and Zisserman, 2003b).

The affine image coordinate frame has as origin the point p, the principal point. However, in practise, and for most image processing applications, the standard that has always been

used, is that the origin is the upper left corner of the view, which corresponds to the lower right corner of the image plane.

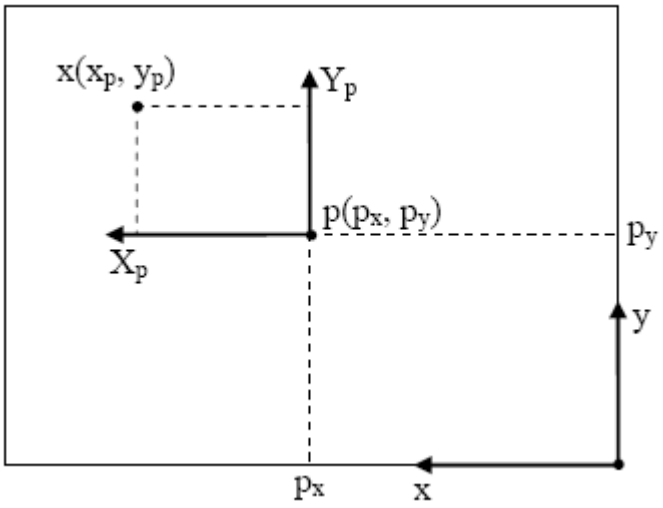


Figure 2.3: Image plane

Introducing this standard and using the parameters defined in Figure 2.3, the projection is expressed as:

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{pmatrix} = \begin{bmatrix} f & & p_x & 0 \\ & f & p_y & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad \textbf{(2.6)}$$

Equation 2.7 has the form:

$$x = K[I|0]X_{cam} \quad \textbf{(2.7)}$$

With

$$K = \begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix}$$

And X_{cam} is the 3D coordinates of the point of the scene in the camera Euclidean coordinate frame.

The matrix K is called the *camera calibration matrix*.

This camera calibration matrix maps a point in the camera Euclidian frame to a point in the image Euclidean coordinate frame.

If m_x and m_y are the number of pixels per unit distance in an image coordinate in the x and y directions and s is the skew parameter and we define $\propto_x = f\, m_x$ and $\propto_y = f\, m_y$, then K, the camera calibration matrix, is expressed as:

$$K = \begin{bmatrix} \propto_x & s & p_x \\ & \propto_y & p_y \\ & & 1 \end{bmatrix} \quad \textbf{(2.8)}$$

Usually, s=0 and the elements of K are called the *intrinsic parameters* of the camera. The camera calibration matrix maps a point in the camera Euclidian coordinate frame to a point in the affine image coordinate frame, however, for most applications, the points of scene are given in the world Euclidian coordinate frame. Therefore, a transformation of the coordinates of any point from a scene from the world coordinate frame to the image camera Euclidian frame is necessary.

This transformation is the sum of rotations and translations. It can be expressed as:

$$X_{cam} = R(X - t) \quad \textbf{(2.9)}$$

Where R is 4 x 4 rotation matrix and t is the 1 x 4 translation vector describing the camera centre relative to the world Euclidian coordinate frame.

The elements of R and t are also known as the six *extrinsic parameters* of the camera [41,43].

2.1.2- Two view geometry:

2.1.2.1- Epipolar geometry:

"The epipolar geometry is the intrinsic projective geometry between two views". (Hartley and Zisserman, 2003d).

It is the geometry dealing with the intersection of the planes including the baseline with the images planes. The baseline is the line joining the two camera centres. It is essential for the search of corresponding points in two stereo images.

In the previous paragraphs the assumption is that the centre of the camera C is located between the camera image plane and the scene. In the following section, for the purpose of illustration, the assumption is that the image plane is placed between the centre of the camera C and the scene. This change does not affect the geometry governing the techniques but facilitates an easier understanding of the principles involved.

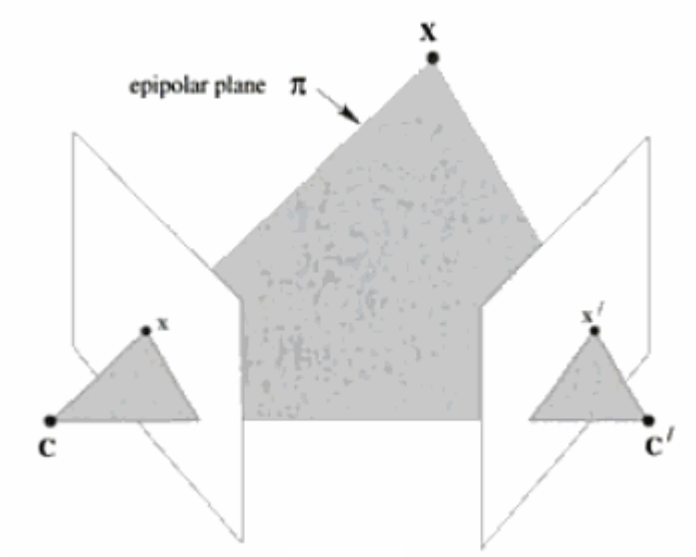


Figure 2.4: Two camera views with image planes in front of camera centres

For given point X from space, x and x' are the image point of X in the two views. C and C' are the centres of the cameras. As Figure 2.4 shows, the point X as well as the cameras centres C and C' and the corresponding images points x and x' are coplanar (π).

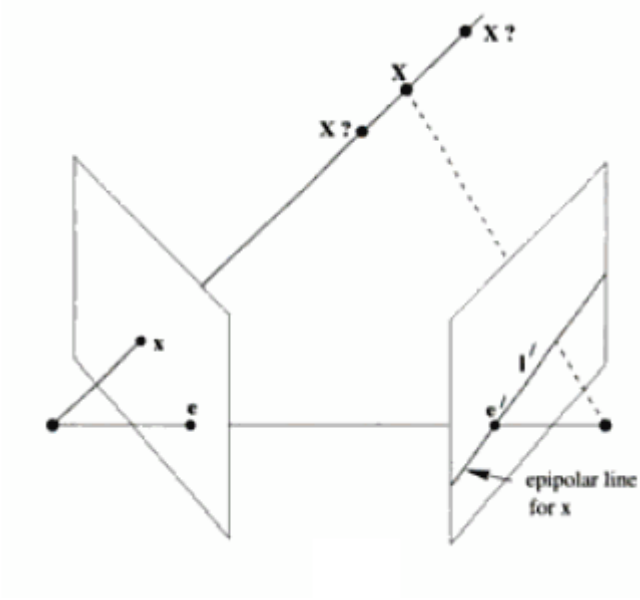


Figure 2.5: Epipolar line representation

Figure 2.5 illustrates how from the image of a point X in one view, there is a corresponding line in the other view. The image of point X on the other view must lie on that line. This line is known as an epipolar line.

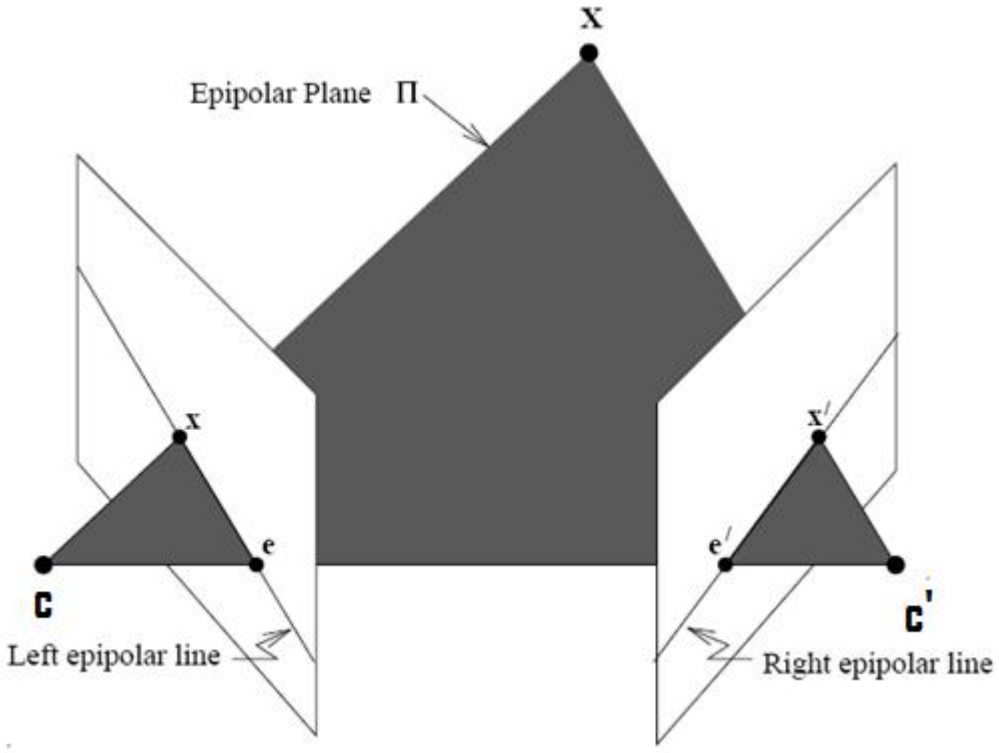


Figure 2.6: Epipolar geometry representation

Figure 2.6 illustrates the entities involved in epipolar geometry; the epipole is the point of intersection of the line joining the camera centres (the baseline) with the image plane. It is also the image point in one view of the camera centre of the other view. An epipolar plane (π) is a plane containing the baseline. Since the epipolar line is the intersection of an epipolar plane with the image plane (Hartley and Zisserman, 2003d), the epipolar line l' is the cross product of the epipole e' and the image point x' (see technical note):

$$l' = e' \times x' \tag{2.10}$$

This is can be written as:

$$l' = [e']_X X' \tag{2.11}$$

where:

$$[e']_X = \begin{bmatrix} 0 & -e'_z & e'_y \\ e'_z & 0 & -e'_x \\ -e'_y & e'_x & 0 \end{bmatrix}$$

Technical note: The cross product of two homogenous 2D vectors results in a line of which both points are an element.

l' is a line, of form $a\,x + b\,y + c = 0$ which is equivalent to $(a, b, c)^T$

the points e' (e'_x, e'_y) and x' (x'_x, x'_y) are on the line l'.

On other hand the product $e' \times x'$ gives:

$$e' \times x' = \begin{vmatrix} i & j & k \\ e'_x & e'_y & 1 \\ x'_x & x'_y & 1 \end{vmatrix} \quad \textbf{(0.1)}$$

$$= i\left(e'_y - x'_y\right) - j(e'_x - x'_x) + k(e'_x x'_y - e'_y x'_x)$$

If the determinant is equal to zero, $i\left(e'_y - x'_y\right) - j(e'_x - x'_x) + k\left(e'_x x'_y - e'_y x'_x\right) = 0$ is of the form $ax + by + c = 0$ which is the equation of a line with $-a/b$ as the value of the slope

Let's verify if points e' and x' are on this line:

$$e'_x\left(e'_y - x'_y\right) - e'_y(e'_x - x'_x) + 1\left(e'_x x'_y - e'_y x'_x\right) = 0 \quad \textbf{(0.2)}$$

$$x'_x\left(e'_y - x'_y\right) - x'_y(e'_x - x'_x) + k\left(e'_x x'_y - e'_y x'_x\right) = 0 \quad \textbf{(0.3)}$$

2.1.2.2- The fundamental matrix:

"The fundamental matrix is the algebraic representation of the epipolar geometry".

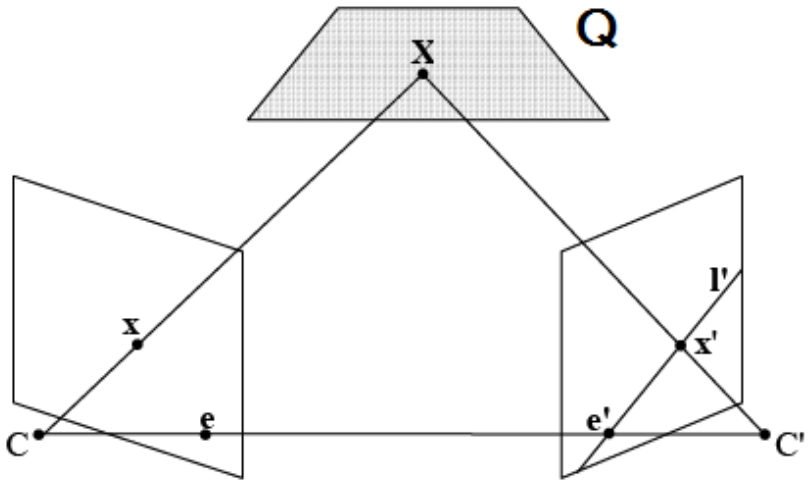


Figure 2.7: Transfer view via plan Q (Hartley and Zisserman, 2003d)

Q is any plane not containing the cameras centres. If X is a point of in this plane, x and x' are respectively the image of X in the left image plane and right image plane. Thus, x and x' are related by a direct linear transformation, known as a homography H_Q. Thus we obtain the identity (Hartley and Zisserman, 2003d):

$$x' = H_Q\, x \tag{2.12}$$

Equation 2.13 will be then:

$$l' = [e']_X\, H_Q\, x \tag{2.13}$$

We define the fundamental matrix F as:

$$F = [e']_X\, H_Q \tag{2.14}$$

As a result:

$$l' = F\, x \tag{2.15}$$

Although the line l' passes through the point x', thus:

$$x'^T\, l' = 0 \tag{2.16}$$

Replacing l' by F x in the equation 2.17 gives:

$$x^T F\, x = 0 \tag{2.17}$$

2.1.2.3- Reconstruction:

The reconstruction is the 3D location of any point X from the stereo images points, also called triangulation.

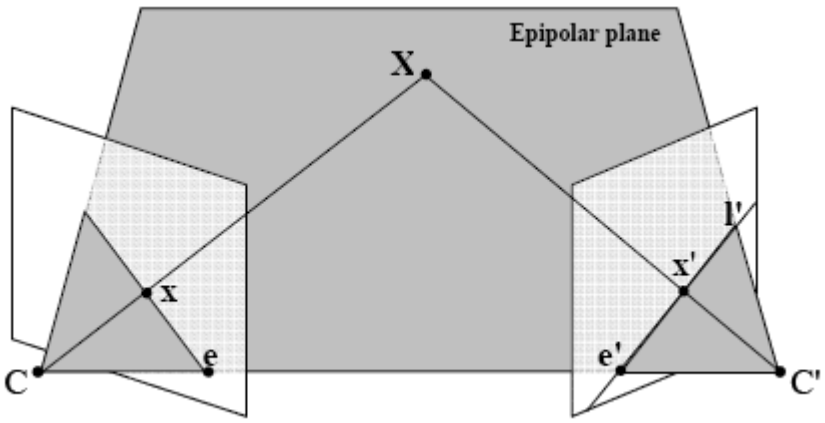


Figure 2.8: The epipolar plane

Once the camera matrices P and P' and the corresponding points x and x' are known, the point X can be triangulated. This is called stereo triangulation; the points x, x' and X are all in the same epipolar plane π. The back projection rays of x and x' through their respective camera centres will therefore have a point of intersection (unless X lies on the baseline). The rays intersect at X (Hartley and Zisserman, 2003a).

An equation A X = 0 can be defined from the equations of the two cameras: x = P X and

x’ = P’ X = P’$_{C2}$ [R|t] X where P’$_{C2}$ is the matrix of camera 2 expressed in the frame related to camera 2 and:

$$x = \begin{matrix} x_x \\ y_y \\ 1 \end{matrix},\ x' = \begin{matrix} x'_x \\ y'_y \\ 1 \end{matrix}\ and\ X = \begin{matrix} X_X \\ Y_Y \\ Z_Z \\ 1 \end{matrix} \quad \textbf{(2.18)}$$

Replacing x and x’ in equation 2.17 gives:

$$(\mathrm{P'X})^T F\ \mathrm{P\,X} = 0 \quad \textbf{(2.19)}$$

It may be shown [41b] that if P is chosen so that P= [I|0] and using equation 2.19 then;

$$\mathrm{P'} = \mathrm{P'_{C2}[R|t]} = [[\mathrm{e'}]\ \mathrm{X\ F|\ e'}] \quad \textbf{(2.20)}$$

The cross product of the point X of the 3D space and its image point x is zero, therefore:

$$x \times (PX) = 0 \quad \textbf{(2.21)}$$

Which results in:

$$x_x\,(p^3_{row}\ X) - (p^1_{row}\ X) = 0 \quad \textbf{(2.22)}$$

$$x_y\,(p^3_{row}\ X) - (p^2_{row}\ X) = 0 \quad \textbf{(2.23)}$$

$$x_x\,(p^2_{row}\ X) - x_y\ (p^1_{row}\ X) = 0 \quad \textbf{(2.24)}$$

Similarly:

$$x' \times (P'X) = 0 \quad \textbf{(2.25)}$$

Which results in:

$$x'_x\left(p'^3_{row}\ X\right) - \left(p'^1_{row}\ X\right) = 0 \quad \textbf{(2.26)}$$

$$x'_y\left(p'^3_{row}\ X\right) - \left(p'^2_{row}\ X\right) = 0 \quad \textbf{(2.27)}$$

$$x'_x\left(p'^2_{row}\ X\right) - x'_y\ (p'^1_{row}\ X) = 0 \quad \textbf{(2.28)}$$

Where p^i_{row} is the i^{th} row of the matrix P and p'^i_{row} is the i^{th} row the matrix P'.

Therefore:

$$A\,X = \begin{bmatrix} x_x\,(p^3_{row}\ ) - (p^1_{row}) \\ x_y\,(p^3_{row}) - (p^2_{row}) \\ x'_x\left(p'^3_{row}\right) - \left(p'^1_{row}\right) \\ x'_y\left(p'^3_{row}\right) - \left(p'^2_{row}\right) \end{bmatrix} \begin{bmatrix} X_X \\ Y_Y \\ Z_Z \\ 1 \end{bmatrix} = 0 \quad \textbf{(2.29)}$$

This algebraic equation contains four linear equations and three unknowns. A non-zero solution for X is determined as the unit singular vector corresponding to the smallest singular value of A (equivalently, the solution is the unit eigenvector of A^T A with least eigenvalue).

The points of the baseline cannot be triangulated. As seen in Figure 2.7, the image points corresponding to the baseline are the epipoles e and e' (Hartley and Zisserman, 2003a; Duffy, 2006).

2.1.3- Camera calibration:

Camera calibration is the process of numerically calculating the parameters values of the camera matrices and the fundamental matrix.

The first paragraph explains the calibration of a single camera and the second paragraph details the calibration of a system of two cameras, also known as stereo calibration.

2.1.3.1- Single camera calibration:

The equation of a single camera, for any point X_i in the world coordinate frame, is:

$$x_i = P\,X_i \qquad \textbf{(2.30)}$$

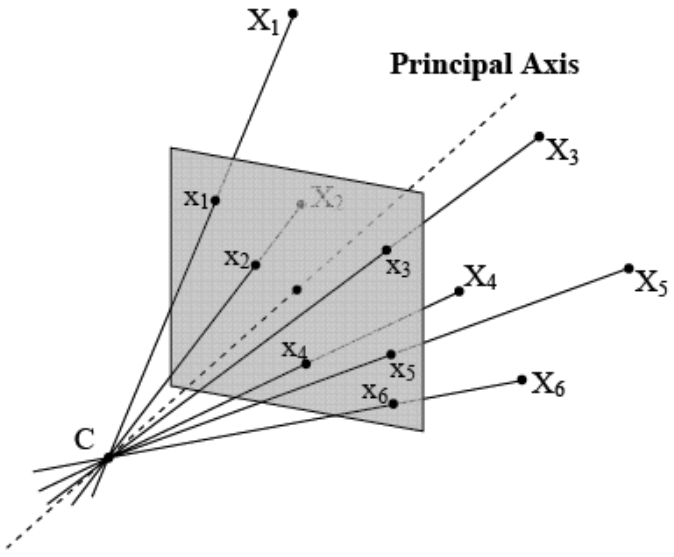


Figure 2.9: Mapping of set of points X_i to x_i (Duffy, 2006)

Each couple X_i and x_i yields the following equation:

$$\begin{bmatrix} su_i \\ sv_i \\ s \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} \quad \textbf{(2.31)}$$

$$x \qquad\qquad P \qquad\qquad X$$

Therefore:

$$\begin{bmatrix} su_i \\ sv_i \\ s \end{bmatrix} = \begin{bmatrix} P_{11}x_i + P_{12}y_i + P_{13}z_i + P_{14} \\ P_{21}x_i + P_{22}y_i + P_{23}z_i + P_{24} \\ P_{31}x_i + P_{32}y_i + P_{33}z_i + P_{34} \end{bmatrix} \quad \textbf{(2.32)}$$

⇨

$$u\,(P_{31}x_i + P_{32}y_i + P_{33}z_i + P_{34}) = P_{11}x_i + P_{12}y_i + P_{13}z_i + P_{14} \quad \textbf{(2.33)}$$

$$v\,(P_{31}x_i + P_{32}y_i + P_{33}z_i + P_{34}) = P_{21}x_i + P_{22}y_i + P_{23}z_i + P_{24} \quad \textbf{(2.34)}$$

The purpose of the calibration process is to solve these equations and determinate the twelve unknowns P_{11} to P_{34}.

If six non-coplanar corresponding points (x_i, X_i) are used then every couple gives two linear equations where X_i (x_i, y_i, z_i) and x (u, v) are known, which gives a total of twelve linear independent equations and twelve unknown values. As a result, these twelve parameters can be solved and thus P determined.

In practice, more than six points are used. Due to errors that will be introduced in any real world measurements, there is not an exact solution of P. However, these errors can be reduced by using a big number of points which will have as the effect of minimising the error to solve the over-determined system.

To generate a set of points X_i, the geometry of the scene must be known. In practice, most researchers use a 'chess board' pattern on a planar surface. The characteristics of this 'chess board', such as grid size and number of squares, must be known and measured accurately. The set of points X_i consist of the corners of the grid. To avoid having coplanar points, several images are taken with the 'chess board' positioned with different angulations with respect to the camera.

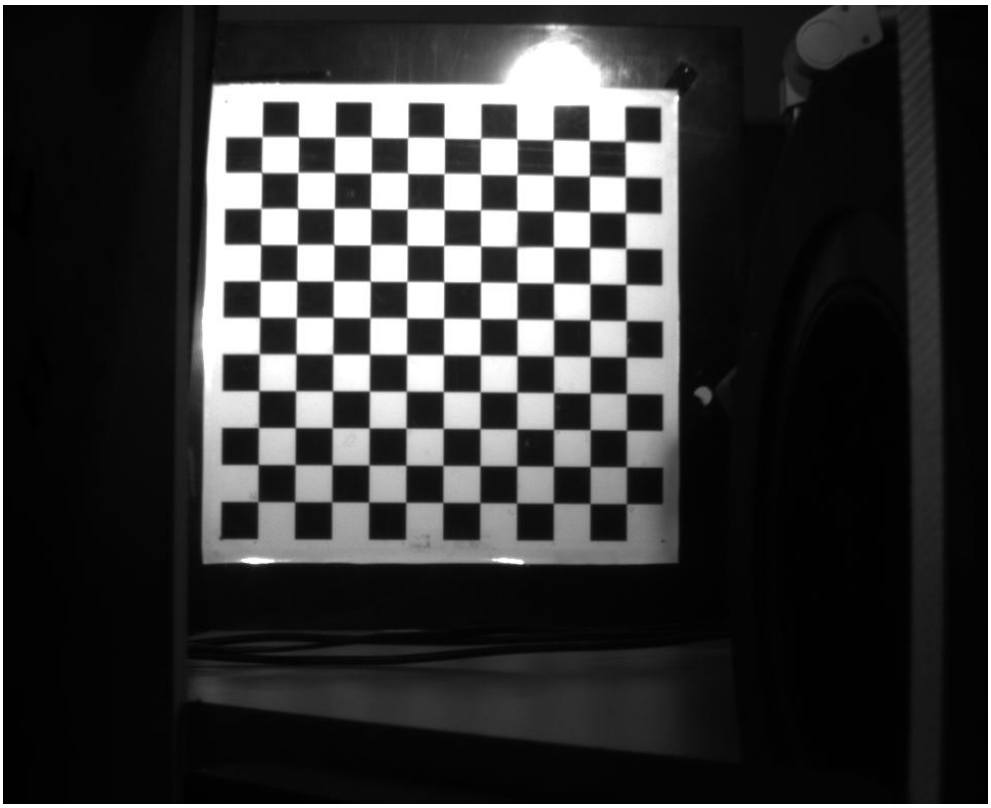


Figure 2.10: Chessboard calibration pattern.

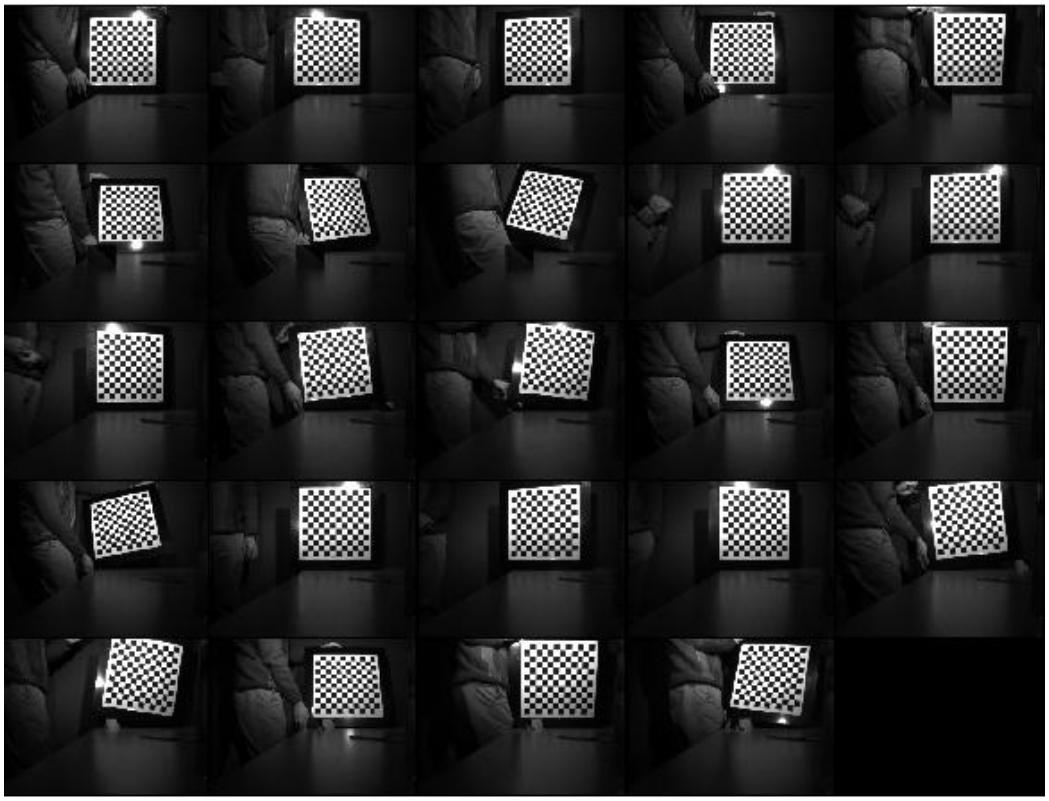


Figure 2.11: Multiple orientation of calibration board.

2.1.3.2- Two cameras calibration:

In the case of two-camera calibration, the process of calibration is defined in the estimation of the fundamental matrix.

Equation 2.33 defines the fundamental matrix:

$$x^T F\, x = 0 \tag{2.35}$$

Where x and x' is a pair of corresponding points in the two cameras.

Expanding equation 2.33 gives:

$$[x' \quad y' \quad 1]\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0 \tag{2.36}$$

⇨

$$[x' \quad y' \quad 1]\begin{bmatrix} xf_{11} + yf_{12} + f_{13} \\ xf_{21} + yf_{22} + f_{23} \\ xf_{31} + yf_{32} + f_{33} \end{bmatrix} = 0 \tag{2.37}$$

Which results in the following equation:

$$x'xf_{11} + x'yf_{12} + x'f_{13} + y'xf_{21} + y'yf_{22} + y'^{f_{23}} + xf_{31} + yf_{32} + f_{33} = 0 \tag{2.38}$$

For a set of n points x_i, x_i', the equation is then written as:

$$\begin{bmatrix} x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ x'_nx_n & x'_ny_n & x'_n & y'_nx_n & y'_ny_n & y'_n & x_n & y_n & 1 \end{bmatrix}\begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = 0 \tag{2.39}$$

A f

To solve for f, a large number of point correspondences are used to give an over-determined system of equations. Due to noise and error in measurements, the point correspondences are not exact so it is very hard to have the exact solution for f. A least-squares solution can be found using numerical methods, where the solution is the unit

eigenvector corresponding to the smallest eigenvalue of $A^T A$ (Hartley and Zisserman, 2003c).

2.2- Image processing techniques:

In developing the routine to identify the teats, some standard image processing techniques have been employed. Two of the important ones, edge detection and image dilation are described here.

2.2.1- Edge detection:

Edge detection is one of the most common techniques used in image analysis. There are probably more algorithms for edge detection than any other technique in the image processing literature (Parker, 1997). The reason is that edge detection is one of the main steps used to determine contours which can then be used for the identification and classification of objects in a scene. Edge detectors can be defined as a set of "very important local image pre-processing methods used to locate changes in the intensity function" (Sonka, Hlavac and Boyle, 1999). The difficulty of finding edges in an image emerges from the concept of an edge itself: for humans, it is easy to identify edges from the shape of an object but for a computer, finding the edge is in itself a step used to identify the object. Another difficulty is the presence of noise, either from digitization of images or optical noise, as well as the presence of different sources of edges in an image such as variable thickness or the variation of shadows depending on the illumination angle of the source lighting (Faugeras, 1993). In this project Canny edge detection, described in the next section, is used.

2.2.1.1- Canny edge detector:

In 1986, John Canny proposed a new approach for edge detection. It is optimal for step edges subject to white Gaussian noise.

The optimality of this approach is related to the satisfaction of three criteria (Faugeras, 1993; Parker, 1997; Sonka, Hlavac and Boyle, 1999):

1. Detection criterion: the edge detector should detect all and only edges, false edges should be missed and all real edges should be detected.
2. Localization criterion: the distance between detected edge and actual edge should be minimized.

3. Response criterion: for a single actual edge, multiple responses should be avoided. This criterion is partially covered by the first criterion so that in the presence of more than one response to an actual edge, false ones should be eliminated. However, this criterion solves the problems of corrupted edges by noise and non-smooth edges.

The Canny Edge detection algorithm uses the rate of change of pixel intensity. An edge would correspond to a sharp increase in pixel intensity and therefore in the region corresponding to an edge the most precise location of the edge will occur when the second derivative of pixel intensity equals zero (corresponding to the maximum rate of change in intensity).

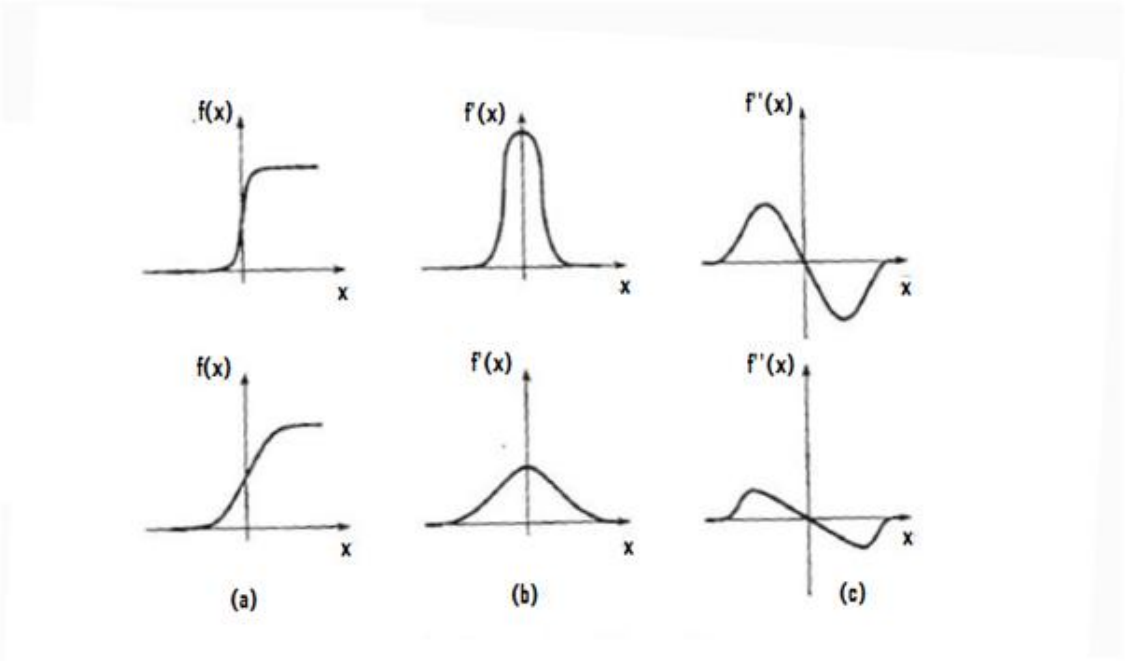


Figure 2.12: 1D illustration of the principle of Canny Edge detection (Sonka, Hlavac and Boyle, 1999)

Figure 2.12 gives a 1d illustration of the principle. Figure 2.12a shows the two different slopes of the step edge profiles of the original image. Figure 2.12b displays the first derivative of the intensity also known as the image function while Figure 2.12c represents the second derivative of the image function. It may be observed how the zero-crossing point of the second derivative is in the same position as the edge in the original image.

The robustness of computing the second derivative is therefore critical and this was considered by Marr-Hildreth (Faugeras, 1993) who smoothed the image initially using a filter in order to reduce the noise, and then computed the second derivative. When choosing the filter, two conditions have to be satisfied: 1- the filter has to be smooth and

band limited; 2- The filter only uses data from nearby pixels (Parker, 1997; Sonka, Hlavac and Boyle, 1999).

A filter that responds to these two criteria is the 2D Gaussian smoothing operator or filter G(x,y), represented by the following equation:

$$G(x,y) = e^{-\frac{x^2+y^2}{2\sigma^2}} \quad \textbf{(2.40)}$$

Where x and y are the image coordinates and σ is the standard deviation of the associated distribution of pixel intensities. Sometimes G is represented with normalizing factor:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad \textbf{(2.41)}$$

Or:

$$\mathrm{G(x,y)} = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{\mathrm{x}^2+\mathrm{y}^2}{2\sigma^2}} \quad \textbf{(2.42)}$$

σ is the only parameter of the Gaussian filter G.

As an example, Table 2.1 contains a convolution mask that is a discrete approximation to a Gaussian function with σ = 1.4. The centre of the mask is passed over each pixel of the image and recalculates the pixel value on the convolution mask. The new intensity value of the active pixel is calculated as the sum of the current intensity value of that pixel and of its neighbouring pixels all multiplied by the corresponding weights contained in the mask. The values are calculated according to Equation 2.42:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad \textbf{(2.43)}$$

0.0105	0.0227	0.0293	0.0227	0.0105
0.0227	0.0488	0.0629	0.0488	0.0227
0.0293	0.0629	0.0812	0.0629	0.0293
0.0227	0.0488	0.0629	0.0488	0.0227
0.0105	0.0227	0.0293	0.0227	0.0105

Table 2.1: Gaussian Low pass filter mask, σ = 1.4

Appling this function to all the pixels blurs the image, reducing the effect of high frequency noise while preserving the image structure. An example of an image containing

Gaussian noise is shown in Figure 2.13. Figure 2.14 is the resulting image after applying a convolution mask such as the one in Table 2.1. The smoothing of the image, and the noise reduction, is seen in the zoomed view (Gonzalez, Woods and Eddins, 2004).

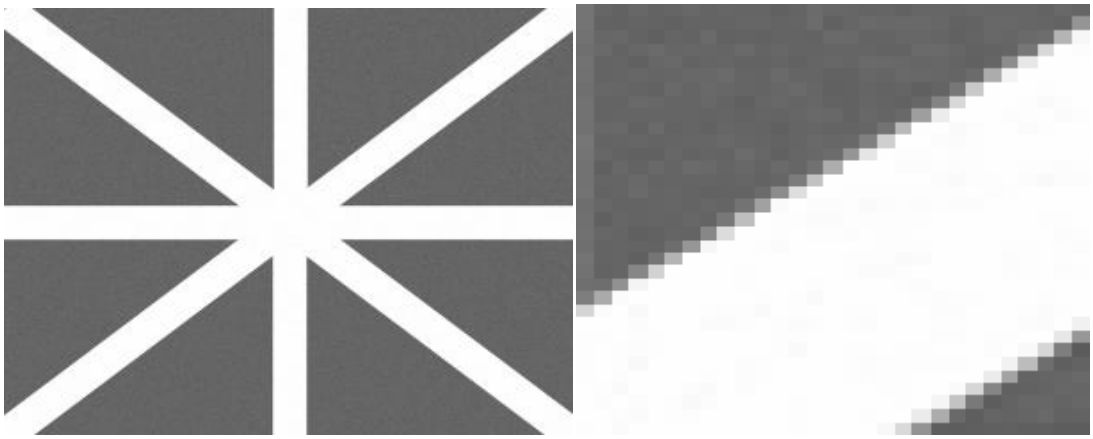


Figure 2.13: Sample image containing Gaussian noise, zoomed view on right.

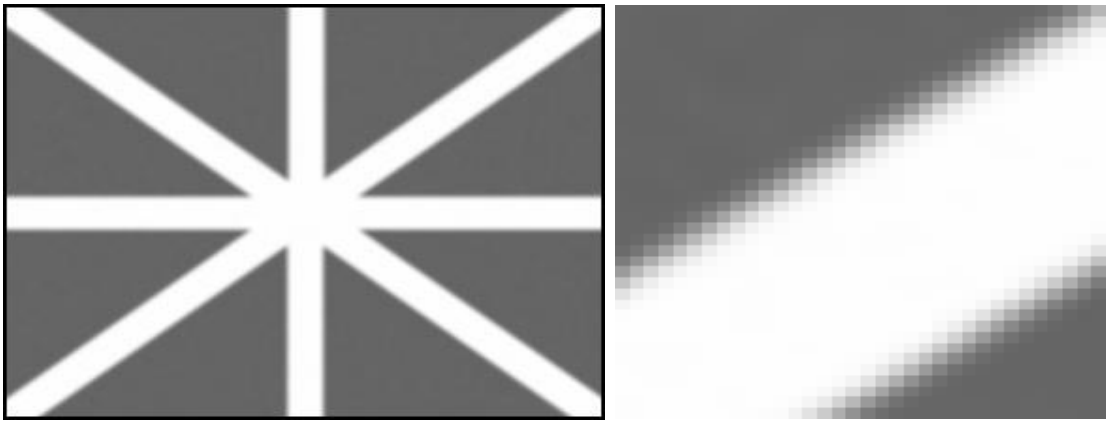


Figure 2.14: Image after applying low pass filter, zoomed view on right (Duffy, 2006)

Using a Gaussian filter is one of the techniques that Canny uses in his edge detection approach. This approach is based on several ideas (Sonka, Hlavac and Boyle, 1999):

1st idea:

After filtering, the edge detector is generalized to 2D so that a step edge is expressed by its location, orientation and magnitude if possible. It can be shown that convolving an image with a 2D Gaussian filter (see previous example) and then differentiating in the direction of gradient generates a simple and effective directional operator. To do this define G_n, the first derivative of G in the direction n:

$$G_n = \frac{\partial G}{\partial n} = n \nabla G \qquad \textbf{(2.44)}$$

The direction n should be perpendicular to the edge. Even if the direction of n is unknown in advance, a robust estimation based on the smoothed gradient direction can be calculated:

$$n = \frac{\nabla(G * f)}{|\nabla(G * f)|} \quad \textbf{(2.45)}$$

where f is the image function. The edge location is then at the local extremum of f convolved with G_n in the direction n:

$$\frac{\partial}{\partial n} G_n * f = 0 \quad \textbf{(2.46)}$$

Substituting equation 2.43 in equation 2.45:

$$\frac{\partial^2}{\partial n^2} G * f = 0 \quad \textbf{(2.47)}$$

Equation 2.45 illustrates how to find local maxima in the direction perpendicular to the edge; it is usually called "non-maxima suppression". As the convolution and derivative are associative operations, the convolution of the image with a symmetric Gaussian filter G can be computed before finding the direction of the second derivative using an estimate of the direction n. The convolution moves from left to right and from top to bottom in the image. It is seen in Figure 2.15 that moving from a grey region to white region results in a high (white) intensity output. Conversely, moving from a white to a grey region results in a low (black) intensity output.

The third characteristic of the edge step, magnitude, is then measured as:

$$|G_n * f| = |\nabla(G * f)| \quad \textbf{(2.48)}$$

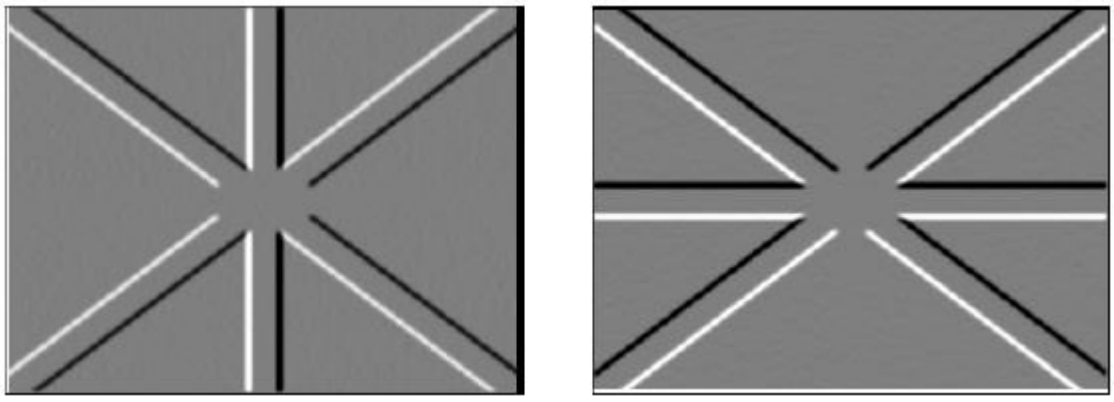


Figure 2.15: Derivate Results: convolution of Gx on left, convolution of Gy on right

2nd idea:

False responses to the single edge caused by noise are a common "streaking" problem in edge detection in general. The output of an edge detector is thresholded to express clearly detected edges. However, breaking up of the edge contour caused by the operator fluctuating above and below the threshold create these streaking problems and therefore a high and low threshold are employed. If any edge response is above the high threshold, these pixels constitute a definite edge output; otherwise, they correspond to noise, unless they are above the low threshold and are connected to pixels with a strong response. The low and high thresholds are set according to an estimated signal-to-noise ratio.

3rd idea:

The correct scale for the filter depends on the objects contained in the image. The solution to this unknown is to use multiple scales and aggregate information from them. Different scales for the Canny detector are represented by different standard deviations σ. There may be several scales that give significant results. In this situation, the smallest scale is chosen and applied to the filter. This will give best localization of the edge.

Canny proposed a feature synthesis approach where all significant edges from the filter with the smallest scale are marked first and the edges of a hypothetical filter with larger σ are synthesized from them. Then the synthesized edge response is compared with the actual edge response for larger σ. Additional edges are marked only if they have a stronger response than the predicted edge. This procedure may be repeated for a sequence of scales, a cumulative edge map being built by adding edges that were not identified at smaller scales.

Algorithm of Canny edge detector:

Combining these ideas leads to the following algorithm (Parker, 1997; Sonka, Hlavac and Boyle, 1999; Gonzalez, Woods and Eddins, 2004):

1. Read in the image to be processed, f.
2. Convolve the image f with a Gaussian G with specified standard deviation σ; the image is smoothed in order to reduce noise.
3. Estimate local edge normal direction n using Equation 2.44 for each pixel of f.
4. Find the location of the edges using Equation 2.46.
5. Compute the magnitude of the edge using Equation 2.47.
6. Threshold edges in the image with hysteresis to eliminate false responses.

For the example treated in this section, the output of this algorithm is shown in figure 2.16 (Duffy, 2006):

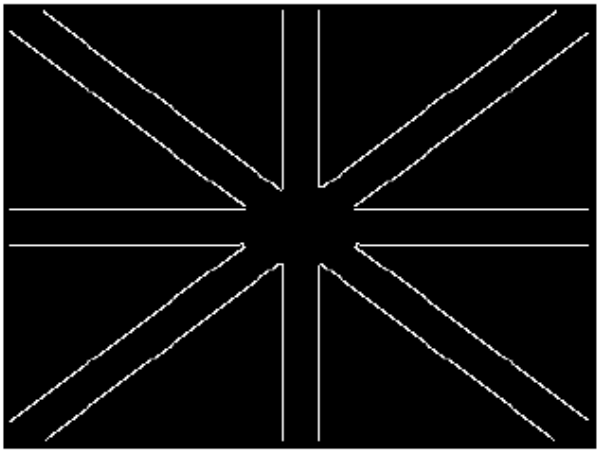


Figure 2.16: Output image of Canny Edge detector

2.2.2- Dilation:

In order to find the locations of the ends of the teats in the optical images captured by the vision system used for this project, some morphological functions, known also as functions of the form and structure of objects (Parker, 1997), have been applied to the images. One of the main steps of the image processing algorithm applied to the optical images is the "dilation".

The dilation has been applied in this algorithm twice; one for a binary image – images with either black or white pixels only-, and one for a grey-scale image – image with "grey" variation pixels.

In this section, an introduction to the theory behind this technique is presented where binary and grey scale images are covered, for more information refer to image processing books (Parker, 1997; Sonka, Hlavac and Boyle, 1999; Gonzalez, Woods and Eddins, 2004).

2.2.2.1- Binary dilation:

Dilation is an operation that amplifies objects in a binary image. The size, nature and shape of this amplification is controlled by a "structuring element". A graphical explanation of the dilation is displayed in Figure 2.17 (Gonzalez, Woods and Eddins, 2004).

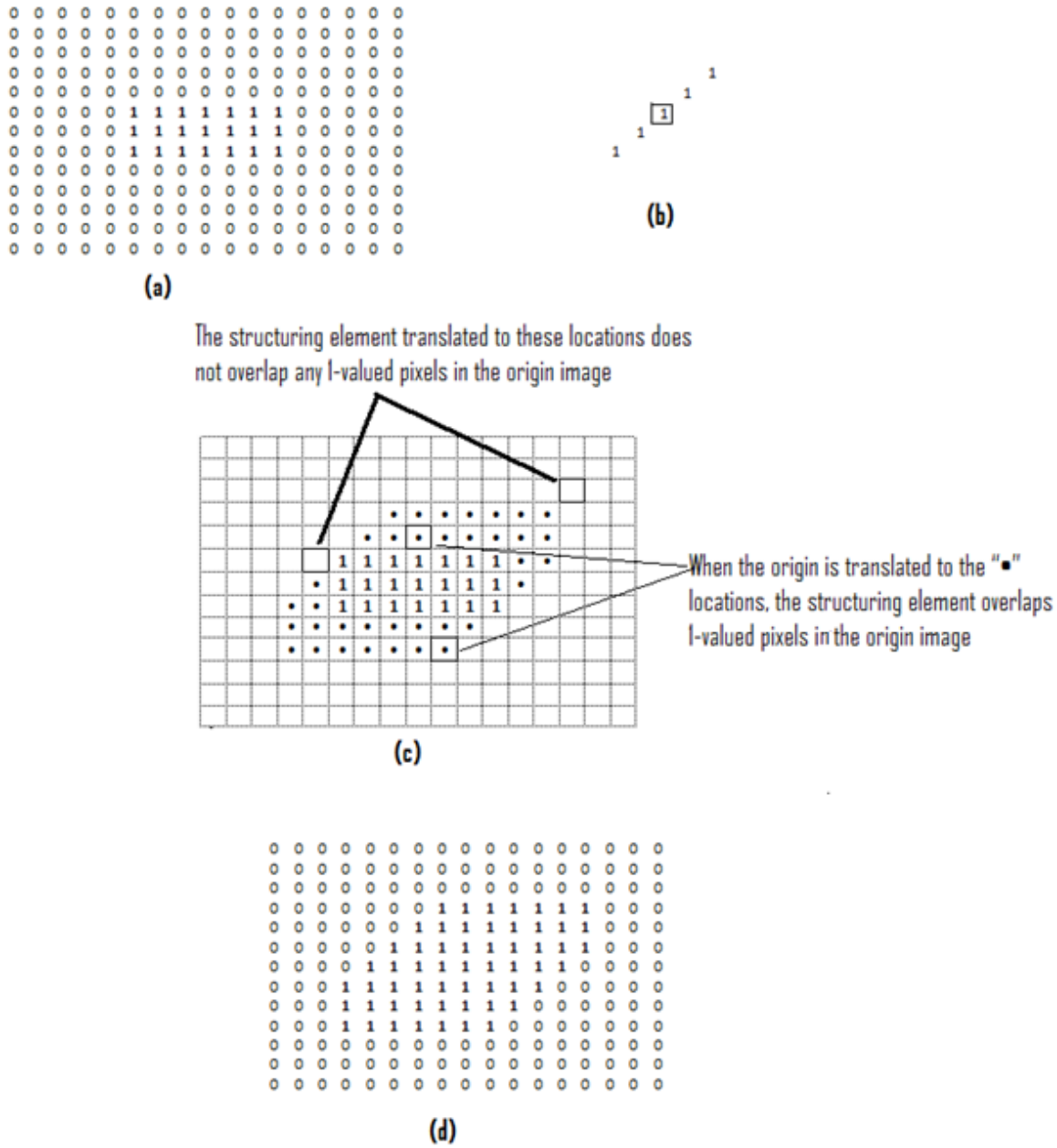


Figure 2.17: Illustration of dilation (Gonzalez, Woods and Eddins, 2004)

The graphics of figure 2.17 represent:

(a) Original image with rectangular object

(b) Structuring element with five pixels arranged in a diagonal line, the origin of the structuring element is shown with dark border.

(c) Structuring element translated to several locations on the image.

(d) Output image

Mathematically, dilation is defined in terms of set operations, so that for a dilation of set A by B (Parker, 1997):

$$A \oplus B = \{c | c = a + b, a \in A, b \in B]$$ **(2.49)**

where A is the image set being dilated and B is the structuring element set.

An image is dilated using a structuring element by superimposing the structuring element on the image. If the image initially holds a 1 at a given pixel, then the output image contains the structuring element with its origin centred on this pixel.

2.2.2.2- Gray-scale dilation:

Although the use of multi grey levels pixels introduces some complication to morphological operation, dilation applied to binary images can easily be extended to gray-scale images. As pixels intensity can take any values in gray-scale images, the definition of dilation as an equation of sets is no longer appropriate and instead a graphical approach is taken. The definition of dilation of grey scale images may vary slightly from one author to another but they all lead to very similar results (Parker, 1997; Sonka, Hlavac and Boyle, 1999).

A mathematical definition of gray-scale dilation can be as follows:

$$(A \oplus S)[i,j] = \max\{A[i-r, j-c] + S[r,c][i-r, j-c] \in A, [r,c] \in S\} \quad \textbf{(2.50)}$$

where S is the structuring element and A is the grey-scale image to be dilated.

This definition can be computed using the following algorithm:

1. Position the origin of the structuring element over the first pixel of the image being dilated.
2. Compute the sum of intensities of each corresponding pair of pixels in the structuring element and the image.
3. Find the maximum value of all of these sums, and set the corresponding pixel in the output image to this value.
4. Repeat this process for each pixel in the image being dilated.

The values of the pixels of the structuring element are also grey levels. They can be also negative. In this case, and because negative values cannot be displayed, two possible solution are used. The negative pixels are set to zero or all pixels level values are shifted in the way that the smallest value becomes zero.

CHAPTER 3

Experimental Set-Up

3.1- Hardware implementation

Once the decision had been taken to combine the stereovision technology with thermal imaging, the next step was the choice of hardware and the physical combination of these two technologies in one multi-modal sensing system.

3.1.1- Presentation of the cameras:

The multimodal sensing system has two optical cameras and one thermal camera. The optical cameras are EC1350 made by PROSLICA (Prosilica, 2005) and the thermal camera is a ThermoVision™ A20 M from FLIR Systems (FLIR, 2004).

The EC1350 cameras are Firewire™ cameras based on the Sony ICX205 AL CCD sensors with a resolution of 1360 x 1024 pixels. The pixels of this sensor are square with a size of 4.65µm (Prosilica, 2005). The lenses used with these cameras are 6mm lenses with manual iris and focus; they are made by NAVITAR Machine Vision (Navitar, 2009). The focal length of the thermal camera is 17mm with a resolution of 160 x 120 pixels.

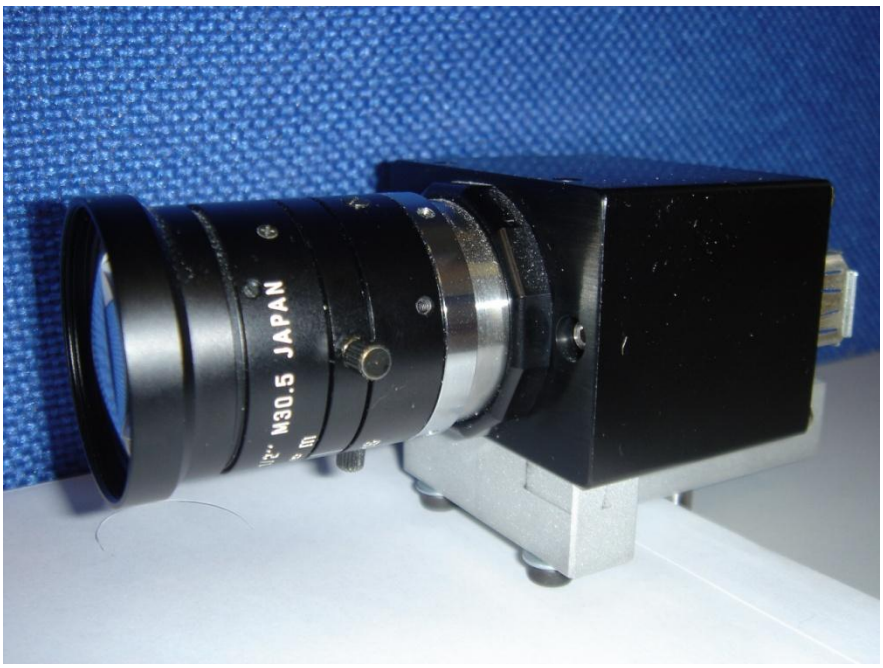


Figure 3.1: Optical camera with lens

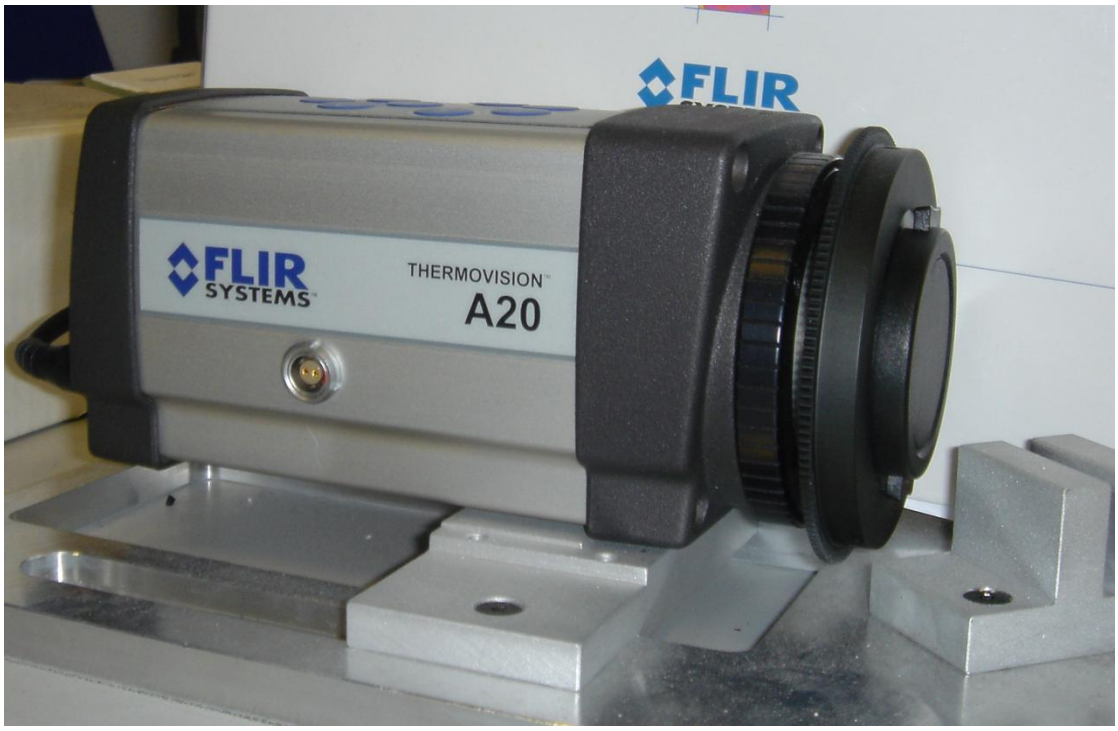


Figure 3.2: Thermal camera

By combining the optical and thermal technologies in one sensing system, the idea is to have the optical cameras capturing the same scene as the thermal camera. Therefore, these cameras need to have the same field of view (FOV) or at least a very similar FOV. However because the optical and thermal cameras available for this project have different focal lengths the FOV of the optical cameras was much larger than that for the thermal camera. For future work the focal length of the optical cameras should be increased to match that of the thermal camera.

In an effort to improve the homography from the thermal to one of the optical cameras, a beam splitter was used. This consists of a sheet of Pilkigton K-glass™ (Pilkington, 2009) which has a special treatment to one side that reflects most of the thermal energy incident on this side. As a result if the glass is placed at an angle of 45deg to the thermal camera and an optical camera with the thermally treated side closest to the thermal camera, the thermal image will be preferentially reflected to the thermal camera while the optical image will be transmitted through to the optical camera see Figure 3.3 (O'Conaire et al., 2005).

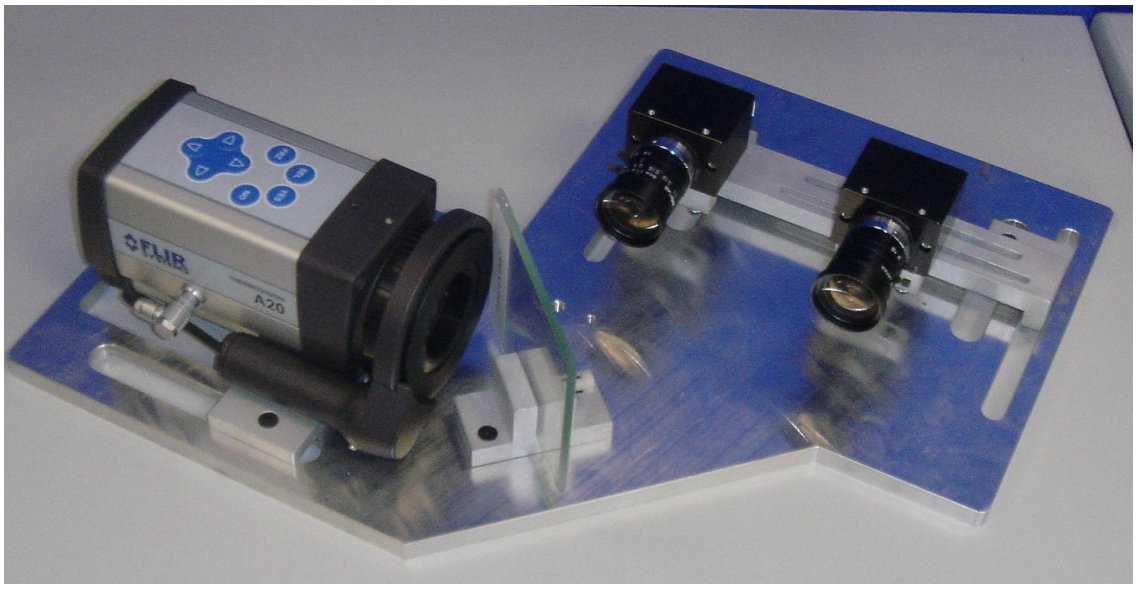


Figure 3.3: Capture rig

3.2- Calibration:

Once the rig had been manufactured and before running any tests, the system must be calibrated. Calibration is the determination of the intrinsic and extrinsic parameters that allow one to equate the coordinate systems of the cameras with one another and a fixed reference frame (see theory, Section 2.1.3).

The calibration is divided in two parts:

1) Calibration of the two optical cameras, known as stereovision calibration

2) Calibration of the thermal camera with the optical cameras to create a thermal/optical homography.

3.2.1- Calibration of the two optical cameras:

Several methods of stereo calibration have been developed. For this project, the calibration has been done using the Caltech's *Camera Calibration Toolbox for Matlab*® (Bouguet, 2008b). Before the stereovision calibration is performed, calibration of the individual cameras is required.

3.2.1.1- Single camera calibration:

The following steps are undertaken during the calibration procedure:

- Acquisition of a patterned array form several different viewpoints. A chess board is used with 24 images recorded.
- From within Matlab, the subroutine calib_gui.m is used to read the images

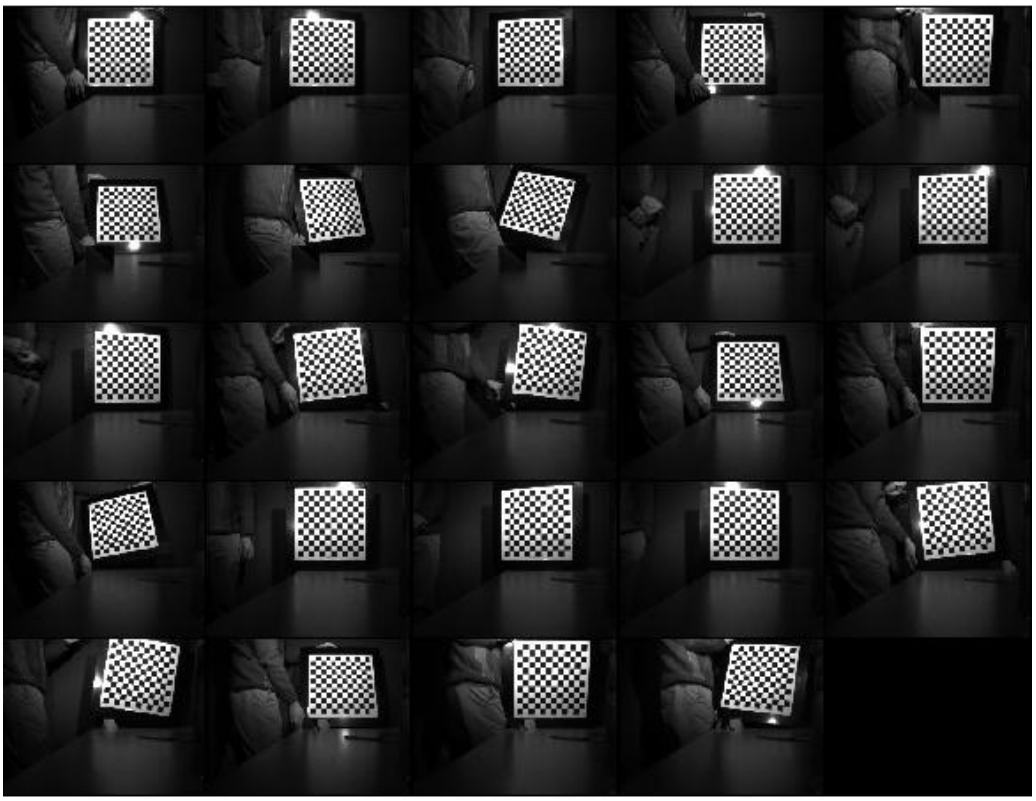


Figure 3.4: A chess board pattern array form several different viewpoints

Figure 3.4 shows 24 images of the chess board pattern array taken from 24 different viewpoints. The purpose of this technique is to have an over determined system of non coplanar points in order to solve the calibration equation (Equation 2.33)

- Following manual identification of the external corners of the patterned array and stipulation of the dimension of individual squares, the algorithm identifies the rows and columns providing lens distortion is not excessive. In the calibration tests performed here, distortion was not found to be a problem.
- The calibration routine for the left camera gave the following results, Table 3.1:

Focal Length:	fc = [1291.0 1295.2]	± [2.7 2.8]
Principal point:	cc = [727.4 528.1]	± [1.2 1.3]
Skew:	alpha_c = [0.0] => Angle of pixel axes = 90.0 ± 0.0 degrees	± [0.0]
Distortion:	kc = [-0.186 0.230 0.0 0.0 0.0]	± [0.003 0.008 0.0 0.0 0.0]
Pixel error:	err = [0.072 0.058]	

Table 3.1: Results of calibration routine for left camera

The results for the right camera using the same method are displayed in Table 3.2:

Focal Length:	fc = [1288.3 1290.6]	± [2.4 2.4]
Principal point:	cc = cc = [701.4 558.7]	± [1.0 1.2]
Skew:	alpha_c = [0.0] => Angle of pixel axes = 90.0 ± 0.0 degrees	± [0.0]
Distortion:	kc = [-0.212 0.301 -0.003 0.0 0.0]	± [0.004 0.012 0.0 0.0 0.0]
Pixel error:	err = [0.066 0.053]	

Table 3.2: Results of calibration routine for right camera

The camera matrix, which was defined previously in the chapter 2 as **K**, Equation 2.7, is stored in the Matlab variable **KK**:

$$KK = \begin{bmatrix} fc(1) & alpha_c * fc(1) & cc(1) \\ 0 & fc(2) & cc(2) \\ 0 & 0 & 1 \end{bmatrix} \quad \textbf{(5.1)}$$

The Caltech's *Camera Calibration Toolbox for Matlab*® includes a function that estimates the distortion and reduces the error. However, for these cameras, the distortion is very low and can be ignored. The error after calibration was also very low.

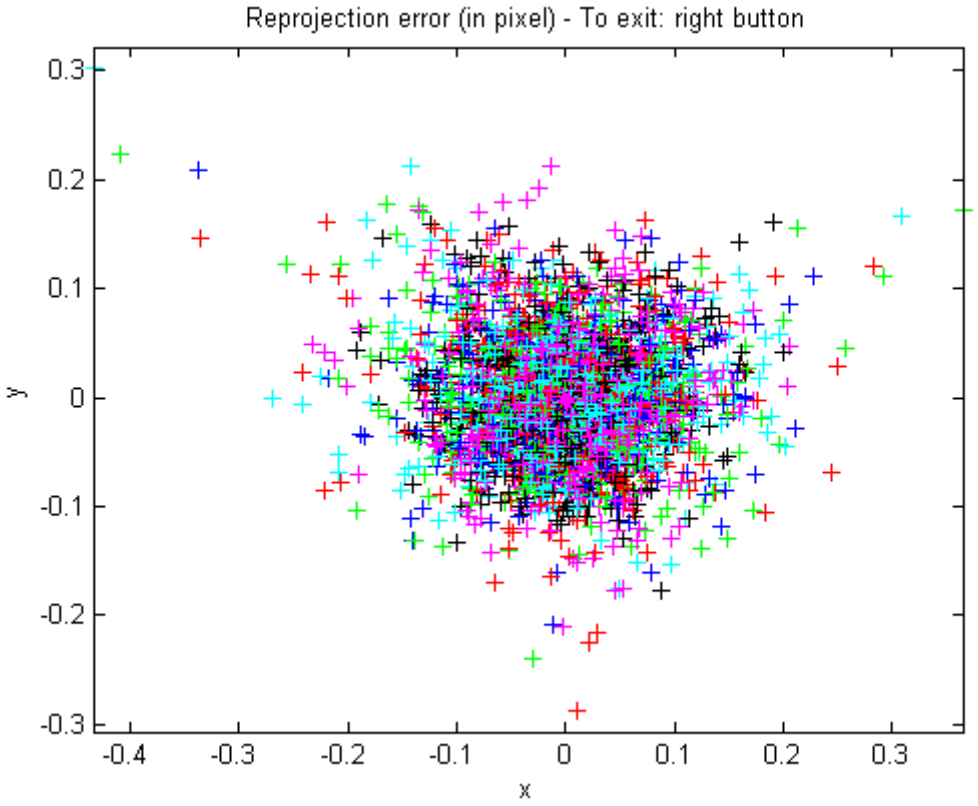


Figure 3.5: Reprojection error

Figure 3.5 represents the error of reprojection after calibration of corners of the calibration pattern array of the left camera. In this figure, the maximum error is of the order of one third of a pixel. To get the best results, it is advisable to run the calibration procedure a few times to optimize the results. For this example, the calibration procedure was run 3 times. In the last one, four gradient descent iterations have been calculated to get the optimized calibration. In fact, for the *Camera Calibration Toolbox for Matlab*®, the optimization is done by iterative gradient descent with an explicit (closed-form) computation of the Jacobian matrix (Bouguet, 2008a). Relative to a given point in the space, the extrinsic parameters can then be determined.
The extrinsic parameters are encoded in the form of a rotation matrix (Rc_ext) and a translation vector (Tc_ext).

3.2.1.2- Stereovision calibration:

Once the optical cameras have been individually calibrated, the stereovision calibration is performed using the Matlab file stereo_gui.m.

The main steps of this calibration are:

- The results of calibration of each camera are loaded.
- Run the main stereovision routine. Results of this calibration consist of the single camera calibration results previously calculated plus the extrinsic parameters:
  Extrinsic parameters (position of right camera relative to left camera):
  Rotation vector om = [-0.0011 -0.0008 0.0085] ± [0.0023 0.0030 0.0003]
  Translation vector: T = [-100.317 0.242 -0.401] ± [0.195 0.218 0.959]

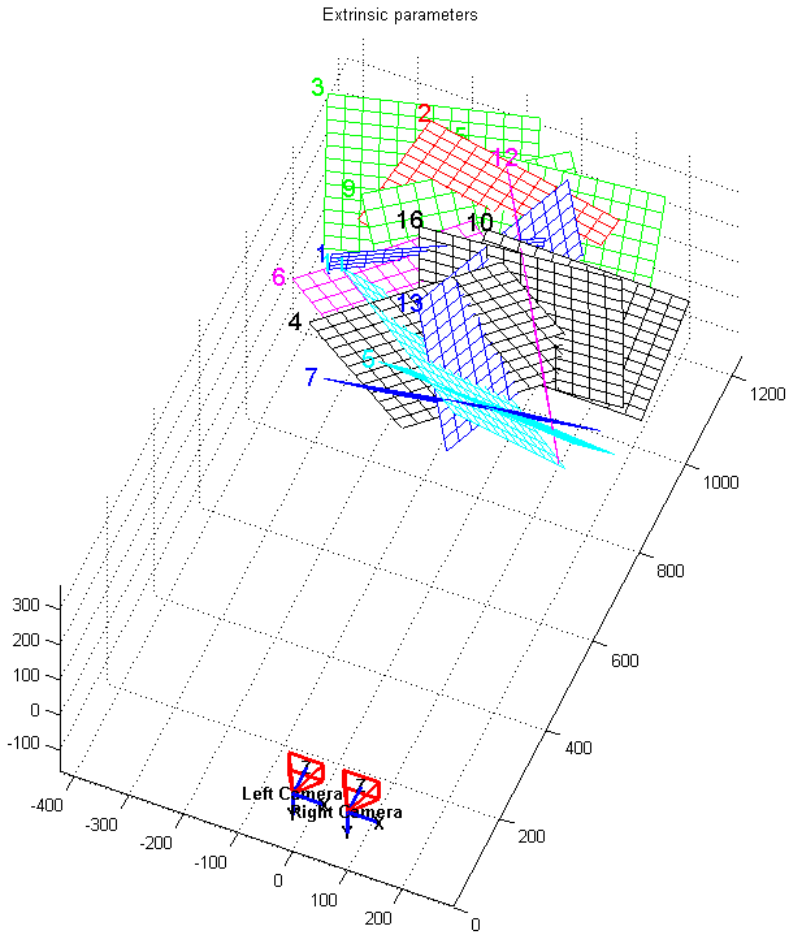


Figure 3.6: Spatial configuration of the two cameras and the calibration planes

Figure 3.6 illustrates the spatial configuration of the two cameras and the 24 pattern calibration planes.

Once the stereovision system is calibrated, triangulation of any point in the field of view of the cameras can be performed. Thus it is possible from the 2-D pixel coordinates of a point in both the left and right camera images to determine the 3-D position coordinates relative to a fixed frame.

In the Caltech *Camera Calibration Toolbox for Matlab*®, this operation is saved in the file stereo_triangulation.m. Some modification has been introduced into the code of this file to make it more suitable in the current application.

The function, stereo_triangulation is called in matlab as:

[XR] = stereo_triangulation(xL,xR,om,T,fc_left,cc_left,kc_left,alpha_c_left,fc_right, cc_right,kc_right,alpha_c_right)

where the arguments in the command line are defined as:

- xL: 2xN matrix of pixel coordinates in the left image
- xR: 2xN matrix of pixel coordinates in the right image
- om,T: rotation vector and translation vector between right and left cameras (output of stereovision calibration)
- fc_left,cc_left,...: intrinsic parameters of the left camera (output of stereovision calibration)
- fc_right,cc_right,...: intrinsic parameters of the right camera (output of stereo calibration)

The output is then given by:

XR: 3xN matrix of coordinates of the points relative to the right camera reference frame

3.2.2- Thermal/optical homographies:

Preliminary testing showed that the reflected image captured by the thermal camera corresponds to the equivalent visual image captured by a virtual optical camera positioned between the two optical cameras. The beam splitter could have been adjusted to match the thermal camera more closely with the adjacent optical camera, but in hindsight since the thermal/optical homography was required from the thermal camera to both of the optical cameras, it was decided to leave the beam splitter alone.

A homography is the planar alignment of data from two different cameras. It is usually needed when it is required to correlate information between two or more cameras viewing a scene from different viewpoints and/or with different operating parameters.

Several methods of data alignment have been reviewed in the literature. Most of these methods have been developed for multimodal vision systems used in medical imagery. In the case of medical applications, a technique for maximising mutual data information is suitable (Viola, 1995; Pluim, Maintz and Viergever, 2003). However, for general applications, as in the current work, a planar homography (Hartley and Zisserman, 2003e) is calculated to align data between the thermal and optical cameras.

To determinate the homography, 40 corresponding points have been selected manually from the thermal image and the optical image. The over determined transformation matrix describing the homography between the thermal and optical images is then solved using a least-squares solution.

Figure 3.7: Optical image used to obtain homography

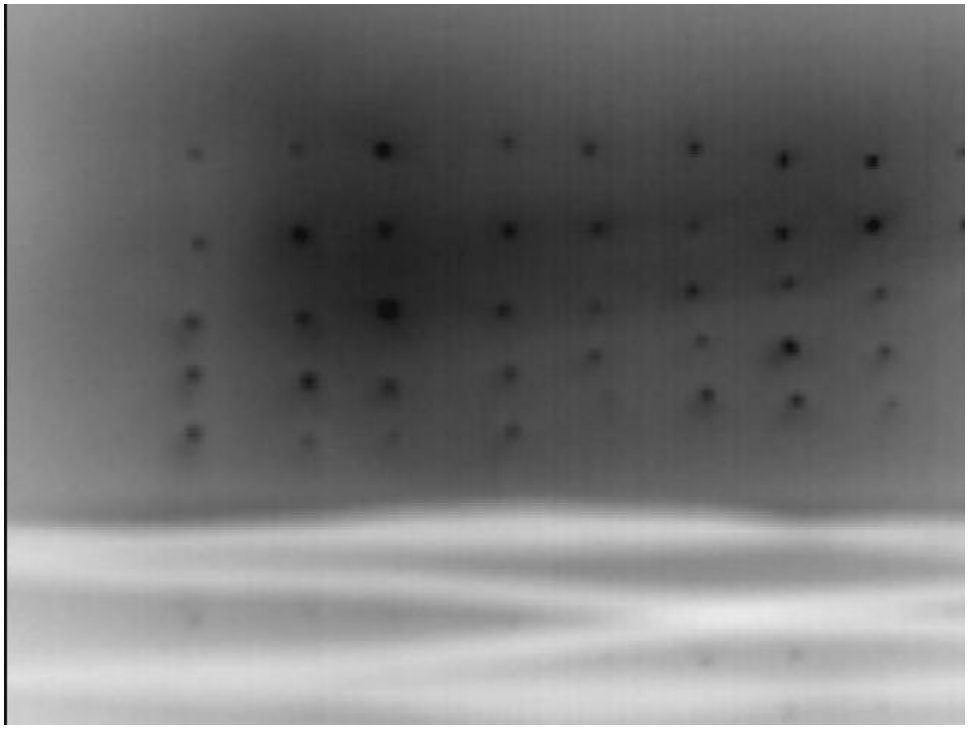


Figure 3.8: Thermal image used to obtain homography

The resultants of homographies are as follows:

Left optical camera/thermal camera homography:

$$\text{left\_homography} = \begin{bmatrix} 0.615 & 0.007 & 672.822 \\ -0.004 & 0.616 & 363.601 \\ 0 & 0 & 1 \end{bmatrix}$$

Right optical camera/thermal camera homography:

$$\text{Right\_homography} = \begin{bmatrix} 0.618 & 0.002 & 536.133 \\ -0.002 & 0.619 & 401.948 \\ 0 & 0 & 1 \end{bmatrix}$$

Preliminary testing has shown that for an object positioned within the volume that would be typical for the udder relative to the camera position, the error in pixel correspondence between the thermal and optical images is no more that 2 pixels in the optical image. For an object within 1700 mm from the camera this corresponds to a distance of 2.51 mm in a plane parallel to the optical image plane.

Using the homography matrix, points determined in the thermal image can be mapped to the optical images.

For example, if $\text{Pixel}_{\text{Thermal}} = \begin{bmatrix} x_{th} \\ y_{th} \\ 1 \end{bmatrix}$ is the pixel coordinates in thermal image

Then, $\text{Pixel}_{\text{optical}}$, which is the pixel coordinates in left optical image, is determinate by the following equation:

$$\text{Pixel}_{\text{optical}} = \text{left}_{\text{homography}} * \text{Pixel}_{\text{Thermal}} \quad \textbf{(5.3)}$$

3.3- Testing Methodology:

3.3.1- Dummy thermal udder:

For testing purposes a phantom thermal udder was manufactured.

Figure 3.9: Thermal dummy teats

The dummy thermal udder consists of a water bath with a temperature control system, a pump and supply and returns lines to the thermal teats, see Figures 3.9 and 3.10.

Figure 3.10: Thermal dummy udder system

The dummy teats have been manufactured from aluminium and painted matt to give them an emissivity characteristic close to that of animal skin.

From field measurements in Teagsac, Moorepark (Moorepark, 2009), the average temperature of the udder and teats of a cow is 32 degrees Celsius. Thus the water bath temperature of water was set in the range of 32-34 degrees Celsius, allowing for a slight drop in temperature as the water is pumped to the phantom teats.

3.3.2. Coordinate reference frames:

During the visit to the farm in Moorpark, one of the criteria examined was the possible position of the teat sensing rig relative to the animal. It was determined that a minimum of 1.5 m from the rear of the animal was required to keep the electronic parts safe and free from the dirt and waste. For the thermal camera used in this project, a distance of 1.5m - 2 m from the scene would be perfect, and would allow the camera to have an optimal field of view (FOV) covering any possible positions of teats see Figure 3.11.

Figure 3.11: FOV of thermal camera for distance of 2m

The distance was finally set to an average of 1.75m for a practical purpose related to the testing environment (length of table).

For ease of testing in the laboratory, the dummy teats were placed on a table upside down, 1.75m from the camera system see Figure 3.12.

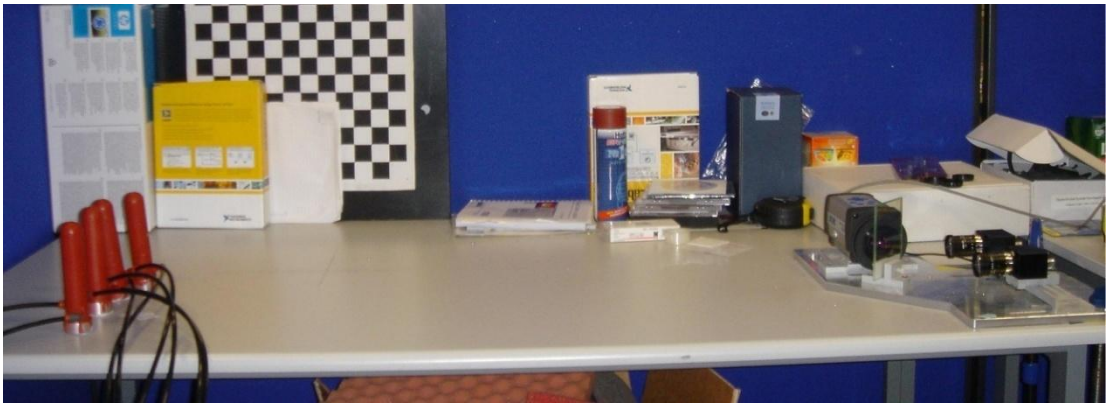


Figure 3.12: Configuration of the cameras- rig and the dummy teats

A teat coordinate frame (which is the world coordinate frame for this work) was defined; the origin O is a point on the table. The surface of the table represents the plane XOY, while the Z direction is orthogonal to the table pointing up see Figure 3.13. It has been set also as the Microscribe coordinate frame system, see figure 3.14.

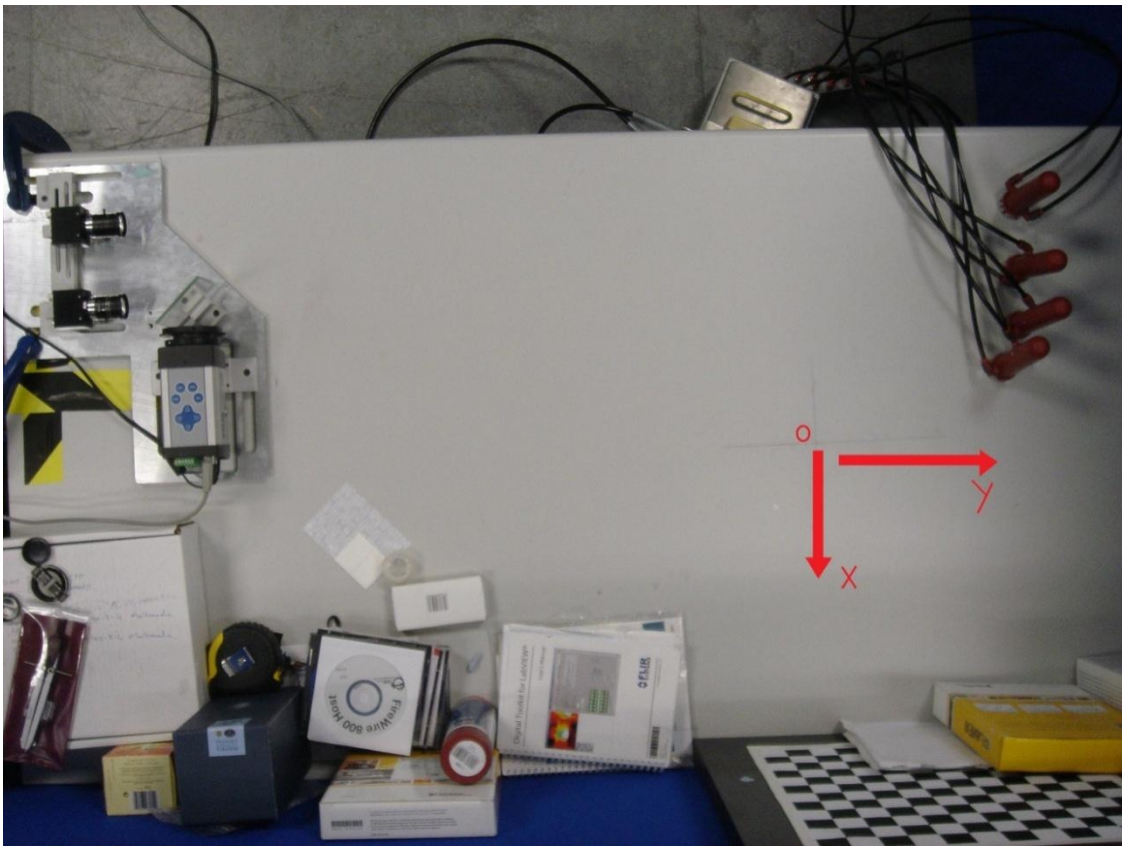


Figure 3.13: Teats coordinates table

This world coordinate frame was also used as the frame of reference for the Microscribe (Figure 3.14), a 3-D digitising device against which the accuracy of the camera system was compared. The Microscribe is a product of Immersion Corporation (Immersion, 2009). The stylus, which is at the end of the arm of the Microscribe, can be freely moved to any point around the base of the Microscribe. Software is used to communicate with the Microscribe via serial communication and read on demand the position of the tip of the stylus. The accuracy of the Microscribe is quoted by the manufacturer to be less than 1mm.

Figure 3.14: Microscribe

The stereovision system has its own coordinate frame, related either to the right or left camera. For this project, the right camera coordinate frame was chosen as the reference frame. The centre of this coordinate frame is situated somewhere in the image plane of the camera, which forms the XOY plane. The X axis points to the right side from the camera, the Y axis points down into the table and the Z axis is perpendicular to the image plane and pointing towards the scene (towards the teats), see Figure 3.15.

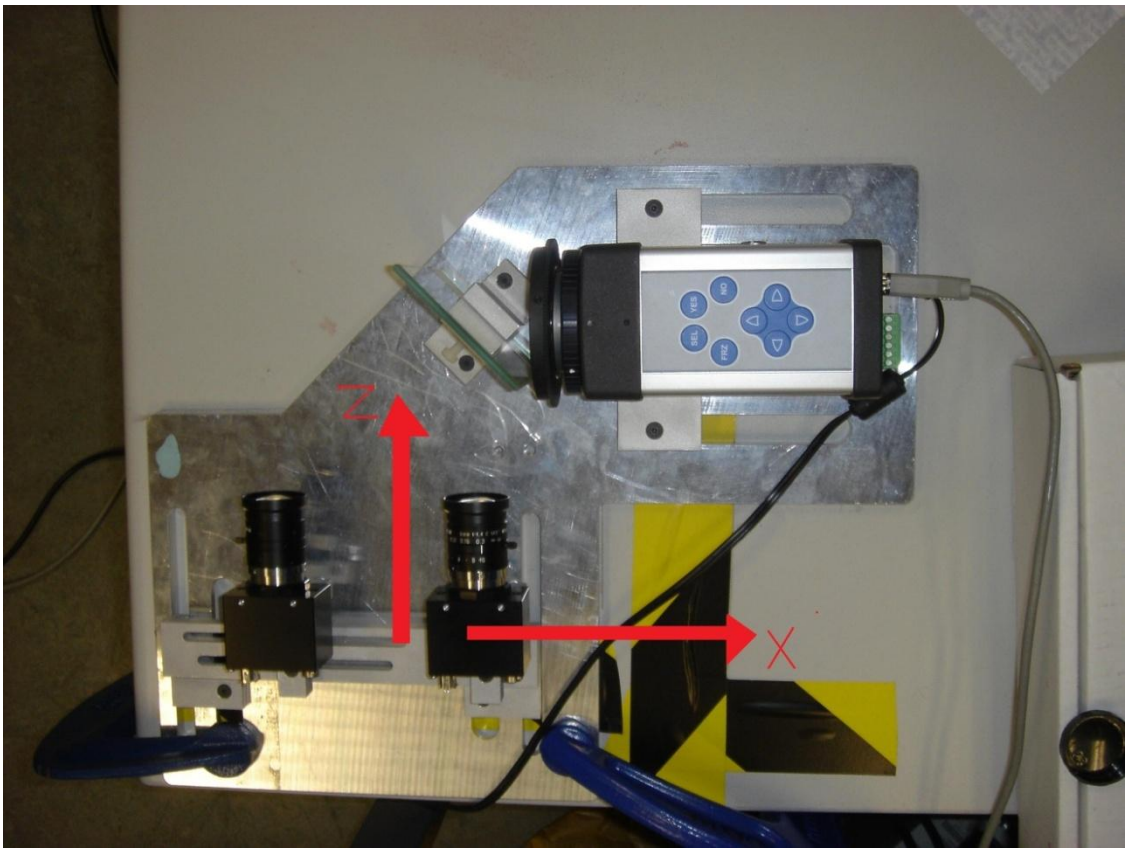


Figure 3.15: Camera coordinate frame

Since the results of the triangulation routine are given relative to the camera coordinate frame it is required to calculate the transformation matrix between the camera coordinate frame and the Microscribe coordinate frame. There are several methods that can be used to calculate transformation matrix between two coordinate frame systems. For this application the method used was a least-square-error technique, it is simple, easy to implement and efficient. Reference points were taken from a chess board pattern, in two different positions covering approximately the space occupied by the teats. 9 points were measured from each position, each point was measured by the Microscribe and the cameras using triangulation.

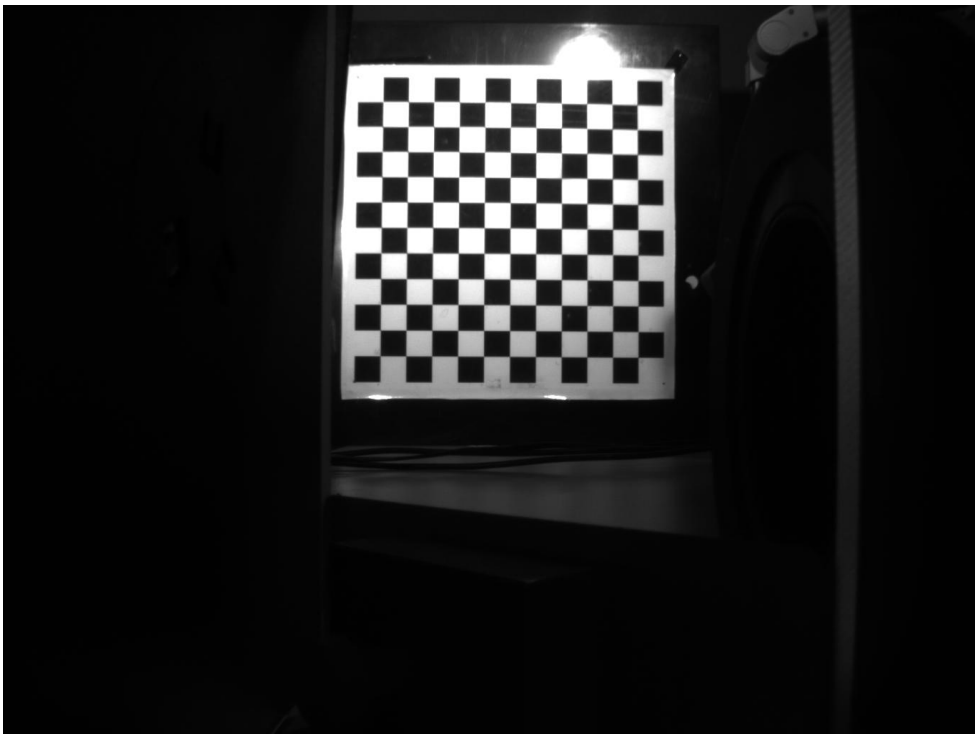


Figure 3.16: Image of chessboard for transformation matrix calculation

The transformation that describes the movement from the camera to the Microscribe frame is a 4x4 matrix containing 12 unknowns. Since right hand orthogonal co-ordinates are used this leaves 11 parameters to be determined. With eighteen measurement points these parameters are over-determined and hence a least squares solution is adopted. Based on the measurement data the following transformation was determined using the '*mrdivide*' function in Matlab:

$$\text{transformation\_matrix} = \begin{bmatrix} 0.99 & 0 & 0.02 & -161.80 \\ -0.02 & -0.01 & 0.92 & -760.40 \\ 0 & -1 & -0.01 & 64.52 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

CHAPTER 4

Teat Identification

Once the system is completely calibrated, the next step is to develop a code to extract information from a captured image thereby identifying teat ends points. An entirely novel algorithm has been developed for this purpose. It is important that the closest pixel corresponding to the same physical location on the teat is determined in the left and right image because it is the difference between the pixel locations that is used to establish the exact 3-D position of the point in the camera frame. As discussed previously in section 1.4.3, spurious objects in the optical images have led to misidentification and hence an initial search for teats is done with a thermal image. The steps of the main algorithm for detection are as follows:

1. Detect the position of end of the teats in the thermal image
2. Apply the homography to the thermal image and locate the approximate position of the teats in the optical images
3. Specify the reduced search region in the optical images for teat identification.
4. Apply an image processing algorithm to find more accurately the end of a teat within the region of interest (ROI).
5. Apply triangulation to end of each teat identified in right and left image; the result is the 3D position of the teat in the camera coordinate frame.
6. Find the 3D position of the teat in the world coordinate frame by applying the transformation matrix between the camera and world frames to the triangulated coordinates

Steps 5 and 6 relate to triangulation and coordinate transformation and have been discussed in Chapter 3, thus steps 1, 2, 3 and 4 will now be investigated.

4.1- Teat recognition- Thermal image

The goal of this routine is to identify and isolate the teats from the background and then to detect the end of these teats. The routines discussed here assume the teats are in full view

in the thermal image; it will not be possible to detect hidden or occluded teats. To find the end of the teats a number of steps are required as follows:

- Image acquisition
- Edge detection
- Find verticals
- Find horizontals
- Find max of verticals
- Find middle of horizontals
- Find end of teats

4.1.1- Image acquisition:

The acquired image by the thermal camera is the thermal energy emitted by the scene and reflected by the beam splitter to the lens of the thermal camera. It is stored in matlab workspace as a variable of type uint8 (unsigned 8 bit integer 0-255) matrix with 480 rows and 640 columns. As a grey scale image; 0 represents black (highest temperature) while 255 represent white (lowest temperature). Intermediate values represent different levels of grey (intermediate temperatures).

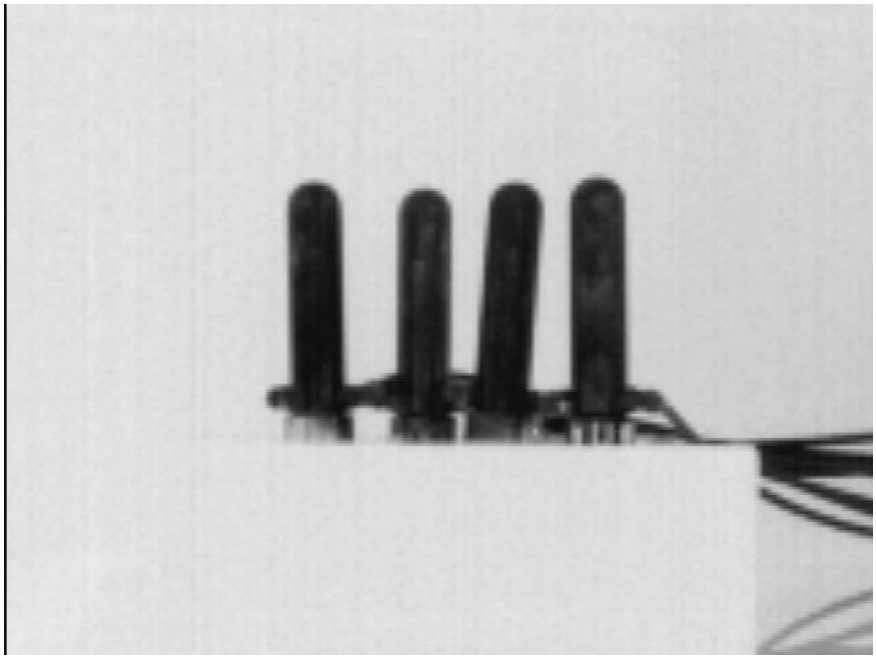


Figure 4.1: Thermal image stored in matlab workspace

Since the thermal image, Figure 4.1, is reflected from the beam splitter it must be reflected again to give the correct representation. The reflection operation consists of reflecting pixels through the centre line of the image.

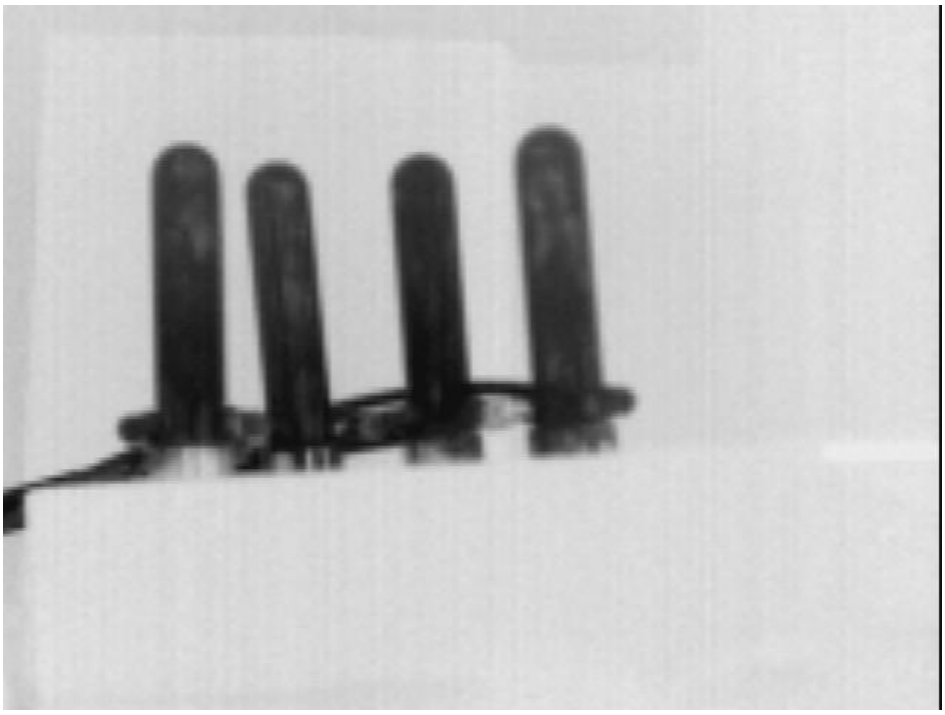


Figure 4.2: Reflected Thermal Image

4.1.2- Edge detection:

The second operation applied to the thermal image is edge detection. Several methods exist for edge detection (see section 2.2.1). During testing several techniques were examined from which the Canny edge detector was found to be the most efficient. The Canny method traces the edges by extracting the local maxima of the gradients in the image. The gradient is calculated following pre-processing of image data with a Gaussian filter. Two thresholds are used; one to detect strong edges and another one to detect weak edges.

In matlab, when using the Canny method for edge detection, there are three variables to set; the two thresholds and sigma which is the standard derivation of the Gaussian filter. The default values for the thresholds were used as determined by the routine. These values are established based on the overall intensity levels in the image.

The third parameter to set is the sigma of the Gaussian filter. Preliminary testing showed that setting this parameter at a high level gave better results. This is due to the fact that the images captured show high differences between the heat intensity of the teats and the background and other objects in scene. For the testing performed in this chapter, sigma was set at 12. For future tests, with real animals, the three parameters can be set to slightly different values depending on the environment.

As shown in Figure 4.3, the result of the edge detection in the thermal image clearly shows the edge of the teats, as well as the tubes of the rig conducting the hot water to the teats.

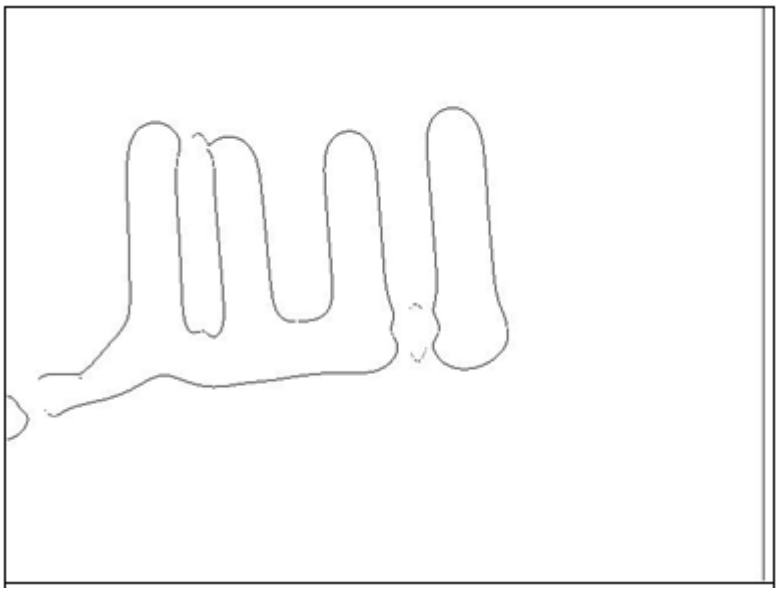


Figure 4.3: Canny edge detection of thermal image (Colours inverted)

In the real world scenario the edge of the udder would also be included in the image and of course the teats would be inverted. Other parts of the animal such as the tail and legs would be removed because they emit less heat than the udder and can therefore be differentiated from the udder region with the appropriate threshold.

The output from the Canny edge detection algorithm is a binary image where edges are set to the value 1 and the rest of the image is set to 0. In this form the image can be further processed to find the teat end points.

After applying these steps to several configurations of the phantom udder, a conclusion emerged; the edge detection routine eliminates almost all exterior objects and the output images of this routine had a certain commonality. Thus an algorithm was devised which first detects all vertical edges that verify certain conditions related to length, location and angle. The second step is to extract all horizontal lines which verify a certain length and are close to the horizontal. Then, from these horizontal lines, select the ones which are located between two of the detected verticals. After that, check that the middles of these horizontal lines are within a prescribed distance vertically and horizontally from the highest points of two the verticals lines. Finally, extract the optimal four points that verify these conditions.

For critical situations, when teats are touching each other or only three teats appear in the thermal image, extra steps in addition to the main routine are required. While this

algorithm may not prove optimal when dealing with live animals, some of the techniques developed may prove useful in the real world application.

4.1.3- Find verticals:

In order to detect the vertical edges, a 2 dimensional statistical filter was used. This requires as inputs, an image, an order and a structuring element (domain). The domain is equivalent to the structuring element used for binary image operations. It is a matrix containing only 1's and 0's; the 1's define the neighbourhood for the filtering operation.
The origin of the domain is placed over the image at a particular pixel. If there is a match between the pixels in the domain and those in the image then the pixel in the image corresponding to the origin of the domain is switched to 1 otherwise zero. This operation can be performed up to an order n where n defines the number of matching pixels that are required between the domain and the image (MathWork Filter, 2009).
The resultant image for the example treated in this chapter has around 50 small segments which represent eight vertical lines, see Figure 4.4.

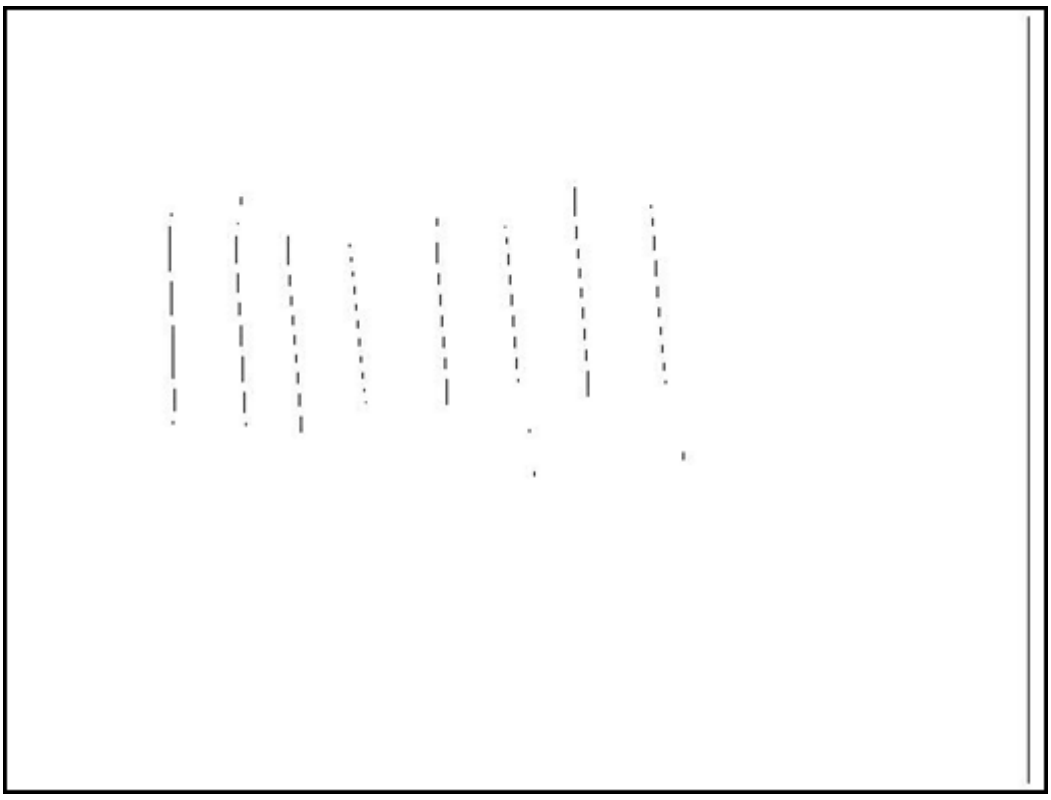


Figure 4.4: Vertical line detection (Colours inverted)

To exclude pixels that are not part of these lines, a series of geometrical conditions were applied as described in the next sections.

4.1.4- Find horizontals

The same method used to find vertical lines was used to detect the horizontal lines. The only change was the 2D domain applied to the filter.

The results of this routine is an image showing around a hundred small horizontal segments. However, it is very clear from the image, that there are four groups of segments that correspond to the ends of the teats. A further routine will assimilate each of these groups of small segments to the teat end position.

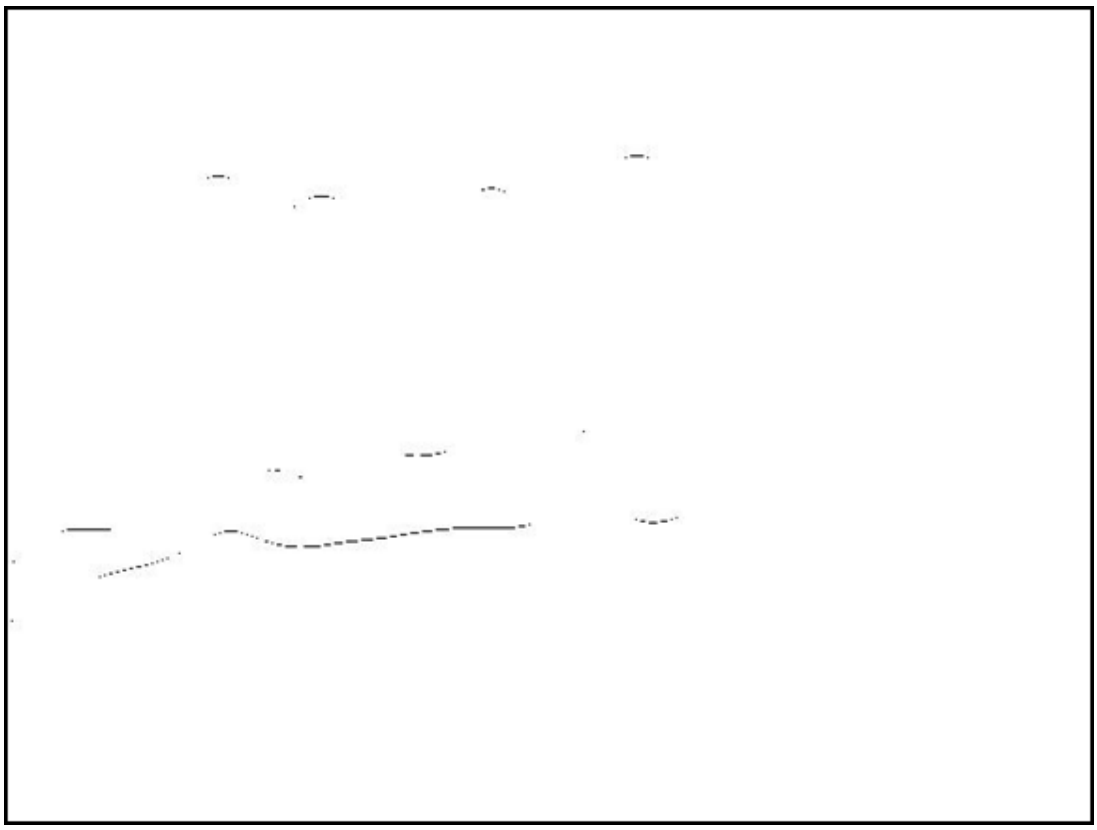


Figure 4.5: Horizontal line detection (Colours inverted)

4.1.5- Find maximum of vertical lines:

As explained previously, any teat end has to verify the following condition; to be within a specific vertical and horizontal distance from the top end of two successive vertical lines. These lines represent the two sides of a teat.

After applying the vertical filter, the resultant image (Figure 4.3) shows a number of vertical lines that are the sum of smaller segments. It is thus required to combine appropriate smaller line segments into larger ones corresponding to the sides of the teat. As a result, an important step is required; get the average vertical position of these lines.

This is done with the following algorithm:

1. From highest row to lowest one, check if any two points have very close horizontal positions. They are then stored with an average horizontal position.
2. Count elements of each position group representative of a vertical line
3. If a group has a sufficient number of elements then it is a vertical line
4. Sort groups according to vertical positions
5. Get the end of vertical lines
6. If there are more than eight lines then extract the optimal eight lines (sides of 4 teats)

The output of this routine is a table with 2 columns and 8 rows, the first column contains the x position of the end of the vertical lines and the second column contains the corresponding y value for each line end. The results are given in the standard image coordinate frame, where the origin is in the top left corner, x increases from left to right and y vertically down from the origin.
For the example above the result is as follows:

	x	y
1st line	104	126
2nd line	148	116
3rd line	182	140
4th line	222	145
5th line	272	129
6th line	316	134
7th line	359	110
8th line	407	121

Table 4.1: End of vertical lines locations

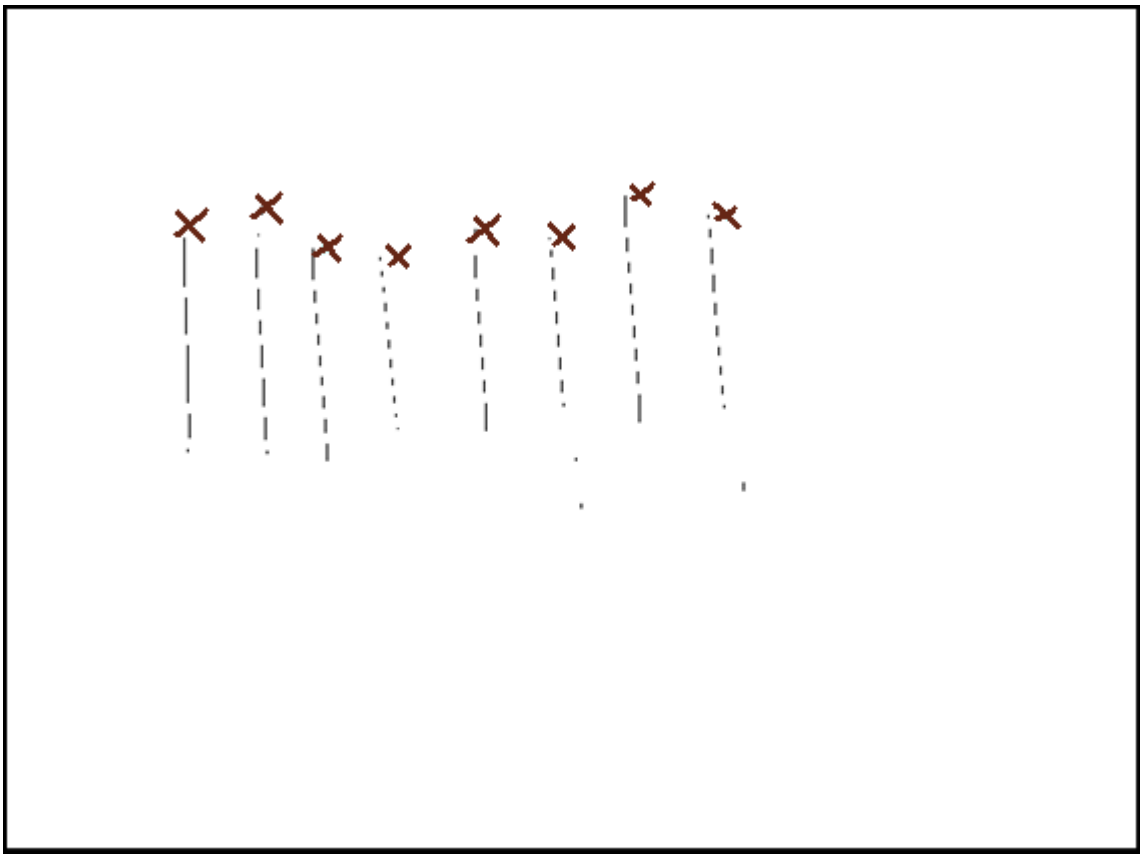


Figure 4.6: Top of vertical lines (Colours inverted)

4.1.6- Find the middle of the horizontal lines:

Before looking for the middle points of the horizontal lines that represent the end of the teats, another step was introduced to improve the performance of this routine. This step consists of calculating the average position of the ends of the vertical lines. Knowing these points helps to eliminate horizontal lines that could not represent end locations of the teats.

Around an average position, a region of interest is selected. Within this region, an algorithm runs to find horizontal lines representing the end of the teats. The strategy used for this task is the same as for identifying the end of the vertical lines.

Therefore, the algorithm is as follows:

1. From left to right, check if any two points have very close vertical positions then they are stored with an average vertical position.
2. Count elements of each horizontal position group
3. Because teats have a typical size, only line segments within a given length range are accepted as potential candidates for an end of teat.

4. Sort groups according to horizontal positions
5. End of teats are then selected as the middle of line segments that are located in a horizontal sense between two of the vertical lines that formed the edge of a teat.
6. If there are more than eight lines, then extract the optimal eight lines. The optimisation is based on the position of each line relative to the previous detected one.

For the example here, the result is as follows:

	x	y
1st point	373	88
2nd point	126	100
3rd point	288	107
4th point	186	112

Table 4.2: Middle of horizontal selection

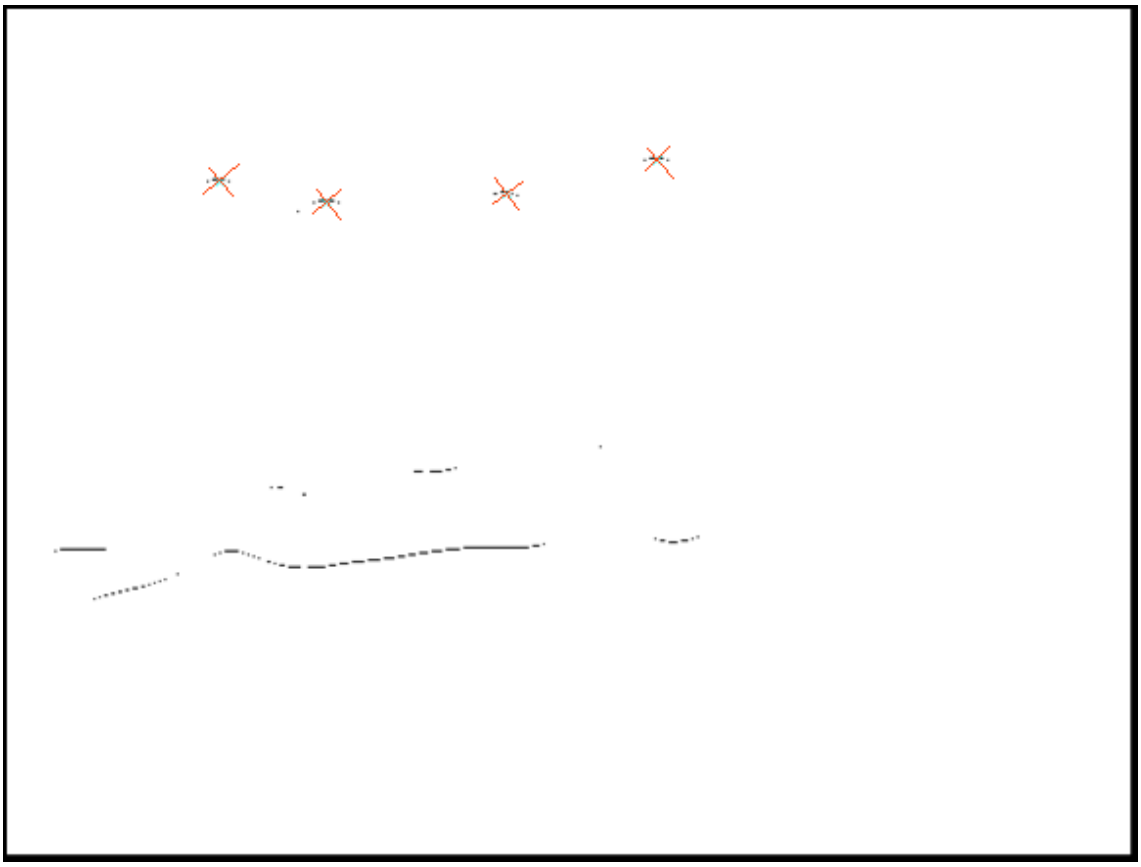


Figure 4.7: Middle of horizontal lines

4.1.7- Find the end of the teats:

This step is the last step of the algorithm applied to the thermal image and consists of looking for the middle of horizontal lines that comply with the condition required for an end of teat: “To be an end of teat, the point has to have x and y coordinates within determined distances from the x,y coordinates of the two ends of the vertical lines” (which will be termed condition C). The algorithm is as follows:

1. Find the average position of potential end of teats.
2. If there are more than four potential points, then eliminate the ones that are not within a determined distance to the average position calculated in 1.
3. Check if the points, left after the selection done in 2, verify the condition of an end of teats, previously set (condition C).
4. Calculate number of hidden teats (4 minus the number of identified teats).

The reliability of the entire algorithm was tested with different teat configurations. These tests are discussed in chapter 5.

For the example treated in this section, the outputs are as follows:

Teat end positions =

126 100

186 112

288 107

373 88

Hidden teats = 0

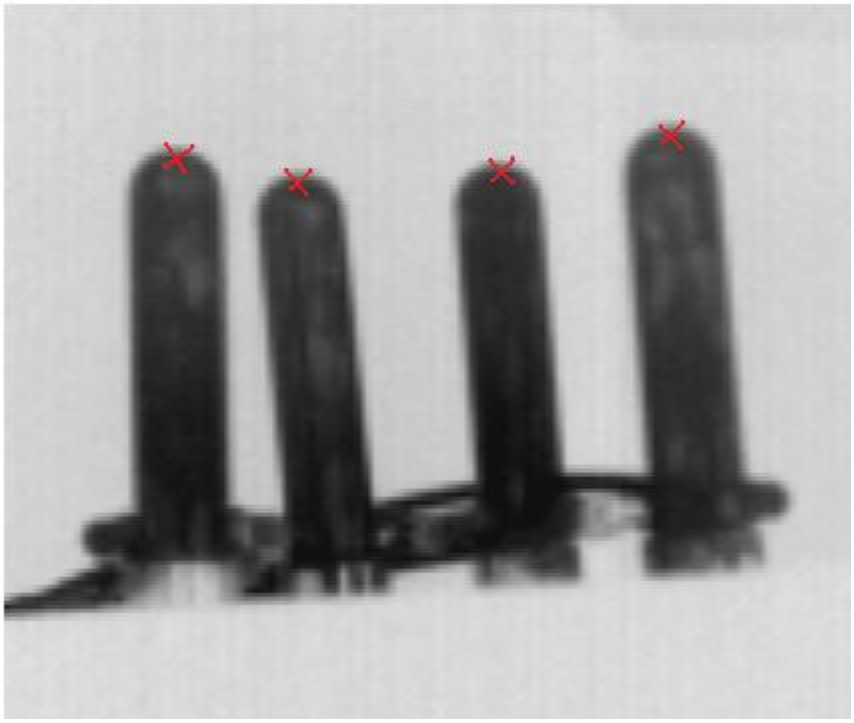


Figure 4.8: Teat ends in the thermal image

4.2- Finding the region of interest:

In this section, a description of the third step of the main algorithm is presented.
The region of interest is specified by establishing a rectangular search region in each of the optical images for four teat positions. The teat end locations are mapped to the optical images from the thermal image using the homographies determined previously (see Section 3.2.2). Around these points the edges of the search rectangle are created based on dimensions ascertained from a number of tests using the phantom teats in various configurations.

For the example here, the results are as follows:

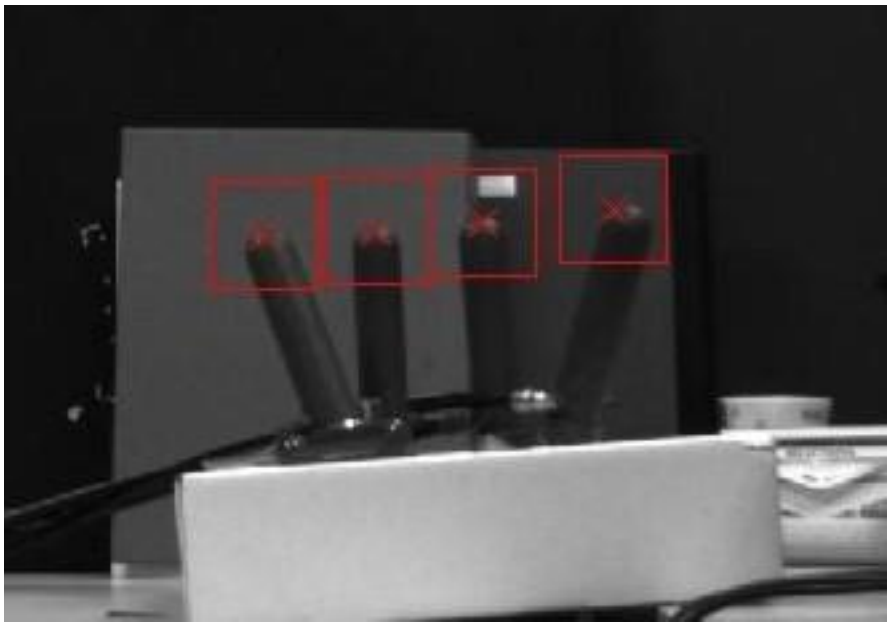


Figure 4.9: Centre of region of interest for the left optical image

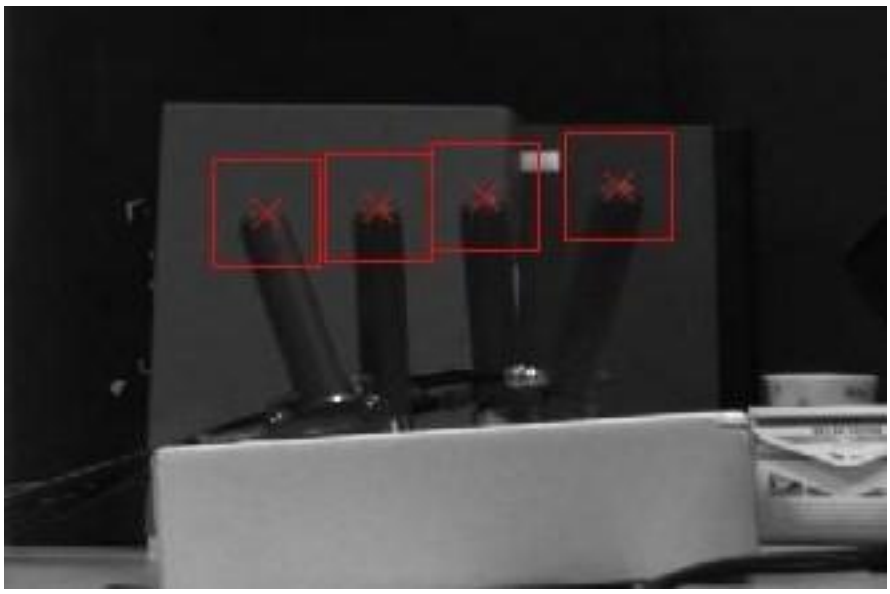


Figure 4.10: Centre of region of interest for the right optical image

From data derived from numerous laboratory tests, the size of the rectangle chosen as the region of interest was set so that it was large enough to include enough pixels to represent the teat of interest and yet not so large as to include any pixel data related to any adjacent teats. This region was determined to be 40 pixels x 40 pixels.

Figure 4.11: Close up of an optical ROI

4.3- Teat Recognition – optical images:

As described previously accurate determination of teat position is obtained using the optical stereo images once the teat region of interest has been established with the thermal camera. In order to do this further image processing is required of the optical data. The first step in the process consists of thresholding the image to distinguish the teats from the background. Following this step it has been determined that the most efficient method of identification is to extract the base of the teat in the cropped region of interest.

Therefore, the algorithm of this routine is as follows:

1. Dilate the image, grey-scale dilation
2. Transform the image to binary image
3. Apply a second dilation, binary dilation
4. Verify the feasibility of the location of the teat
5. Apply mathematical location of the teat

The previous algorithm is applied to each of the four teats in the right and left images.

4.3.1- Grey-scale dilation:

Dilation is one of the most used techniques in image processing. It is a morphological function. More details on this technique are given in section 2.2.2.
This step was performed by applying the 'imdilate' function of Matlab to the input image (MathWork Dilate, 2009). This function requires as input the image and a structuring element object, or an array of structuring element objects. After some preliminary testing, the structural element chosen is a table with a rectangular base and very flat pyramidal shape header, see figure 4.12.

0	0	1	1	1	1	1	1	0	0
0	1	1	1	1	1	1	1	1	0
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1

Figure 4.12: The structural element of the grey-scale dilation

The result of this step is a dilated image where objects with a similar shape to the chosen structural element are enhanced to match that predefined structure. For the example treated here, the base of the teat and also a light reflection on the teat have been enhanced.

Figure 4.13: Result of first dilation

4.3.2- Transform the image to binary image:

The second step of the image processing algorithms is to convert the resultant image into a binary image. The output image of the previous step is a grey scale image, working with a binary image simplifies the classification of the elements in the image and hence feature recognition. This transformation has been realised using the Matlab function ‘im2bw’. This function replaces all pixels in the input image with luminance greater than a given level with the value 1 (white) and replaces all other pixels with the value 0 (black).

Figure 4.14: Resultant binary image

4.3.3- Binary dilation:

In order to highlight the base of the teat, a second dilation has been applied to the resultant image. As in the previous dilation, the function 'imdilate' was used. However, the structural element used in this step is a rectangle 10 by 7 pixels in size. . Figure 4.15 shows the result of this dilation applied to Figure 4.14. In this image, the base of the teat and a rectangular area in the right side of the image have been highlighted. This area has been highlighted because of lack of contrast between this area and the teat. Also, the base of the teat has been widened.

Figure 4.15: Result of the second dilation

4.3.4- Verify the feasibility of the location of the teat:

The identification of the teats with the optical cameras increases the accuracy of the location of the teats. However, in some cases the environment is quite dark or the background is not neutral, and has for example the same colour as the teats. In these situations it is not possible to extract the end of the teat in the optical images.

For example the following image, Figure 4.16, is the output from the second dilation step when there is a non-neutral background:

Figure 4.16: A non relevant result of the second dilation

In this case there are a large number of dark pixels and it becomes impossible to locate the end of the teat and therefore no refinement is performed. For this reason a verification step has been introduced after the second dilation; this step consists of counting the number of black pixels in the teat image. In such cases, relying on the information extracted from the thermal imaging is more reliable.

4.3.5- Mathematical location of the teat:

When searching for the end of the teat, the image (such as that in Figure 4.15) is examined from the top down on a row by row basis until a continuous dark line is found that has a length within a given criterion. The middle of this dark line is obtained and the teat end is established as a point above at a distance equal to half the line length.

CHAPTER 5

Results and Discussion

The viability of the teat sensing system will depend on how accurately it can determine teat position and how robustly it can identify teats from background objects in the field of view. To give an idea of how accurate the current hardware can be, measurements are made of known teat positions after manually identifying the teat locations in the images. This technique avoids the problem of incorrect teat identification, a problem related to image processing rather than the potential accuracy of the hardware. The effect that incorrect teat identification can have will be assessed by purposely picking teat end positions that are displaced from the optimum identifiable in the image. Knowing the errors induced will give some indication of the effect of teat misidentification on positional accuracy.

In the second part of this chapter the teat identification algorithm is assessed in non-ideal conditions. Various scenarios are considered such as partial occlusion, thermal interference etc. and the ability of the algorithm to correctly select the ends of the teats is discussed.

5.1- Accuracy of stereo triangulation:

5.1.1- Test 1:

Once the calibration is done, the next step was to verify the accuracy of the calibrated system by estimating the position of known points. This was achieved by placing four teats in the optical cameras field of view and manually selecting the end positions in the left and right optical camera images and then introducing their coordinates into the triangulation algorithm. The teats were positioned in a standard configuration with neutral background and were clearly seen by the two cameras (Figure 5.1 and 5.2). The aperture of the lens of each camera was reduced to the minimum possible in order to be as close as possible to the theoretical model of a pin-hole camera (Chapter 2). However, when the aperture is reduced, the scene becomes very dark and needs strong lighting. This setting of reduced aperture and strong lighting were used only during the testing of the accuracy

of the triangulation routine. By reducing the aperture the depth of focus is increased improving edge definition in the images. Later, in the section 6.2, some suggestions to improve the limit of the hardware without reducing the aperture are discussed.

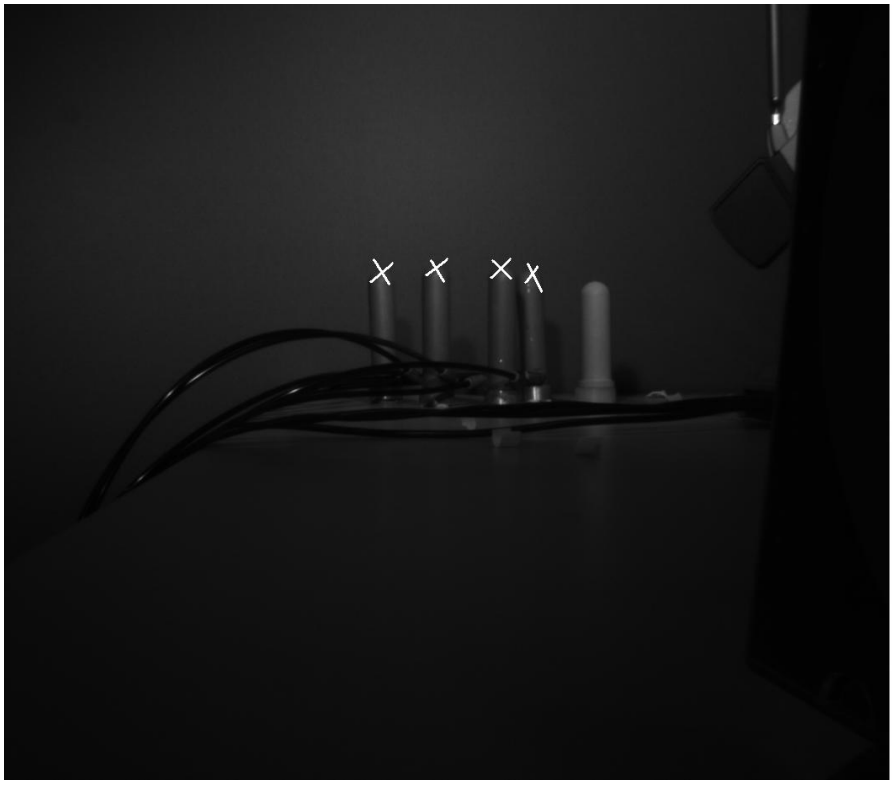


Figure 5.1: Left image of four teats manually selected and marked

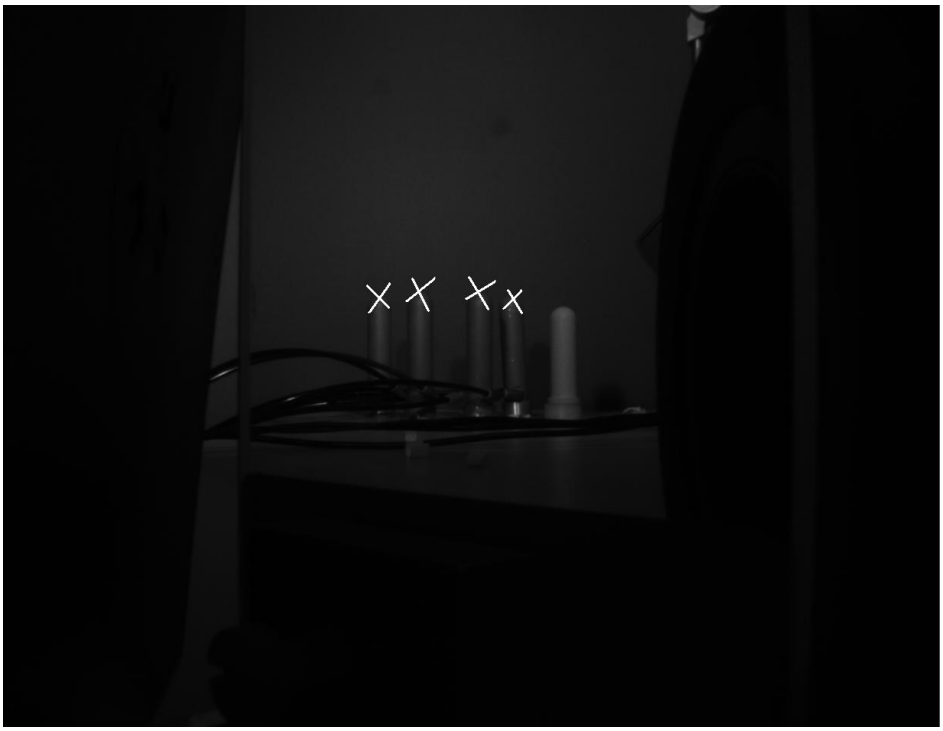


Figure 5.2: Left image of four teats manually selected and marked

The teat end locations were also measured with the microscribe. The table below relates the teat end positions determined with the microscribe and by stereo triangulation.

	X	Y	Z
Teat 1	-110.2	284.2	134.2
Teat 2	-161.9	192.6	137.6
Teat 3	-224.9	199.4	136.7
Teat 4	-260.6	267.1	136.3

Table 5.1: Teat end positions measured with microscribe (mm)

	X	Y	Z
Teat 1	-109.8	285.7	135.5
Teat 2	-161.3	191.7	137.8
Teat 3	-224.2	197.4	137.4
Teat 4	-260.5	269.9	137.4

Table 5.2: Teat end positions calculated by stereo triangulation (mm)

The coordinates are described relative to the microscribe frame. This required a further transformation of the triangulated points from the camera reference frame to the microscibe frame using the transform described in Section 3.3.2. From the data it is apparent that the disparity between the two sets of data is of the order of 1 mm. The accuracy of the microscribe system is quoted as 0.1mm (Immersion, 2009) and therefore it can be assumed that the optical triangulation method can provide an accuracy of 1mm which is well within the required accuracy of ±5mm. The inaccuracy is a result of several factors including errors during calibration, optical effects, image resolution related to pixel density and also operator error when using the microscibe.

5.1.2- Test 2:

The results obtained in Table 5.2 rely on manual identification of corresponding pixels in the left and right camera images. In practical circumstances this will be done automatically using information from the thermal image and signal processing techniques. In this case, these points will be identified with some error.

In order to check the impact of such errors on the triangulation routine, a second test was performed with the same teat positions as test 1 except during the manual selection of the teat ends an error was purposely introduced by offsetting some of the readings as follows:

- Offset the position of teat 1 by 1 pixel in right image and 1 pixel in left image.
- Offset the position of teat 2 by 2 pixels in right.
- Offset the position of teat 3 by 2 pixels in left image.
- Offset the position of teat 4 by 1 pixel in left image.

In this case the triangulation gave the following results:

	X	Y	Z
Teat 1	-105.1	290.8	134.9
Teat 2	-152.8	201.8	137.6
Teat 3	-215.9	206.9	136.9
Teat 4	-255.2	275.2	136.3

Table 5.3: Teat end positions calculated by stereo triangulation after offset (mm)

The errors between the measurements of table 5.3 and 5.1 are as follows:

	X	Y	Z
Teat 1	5.1	6.6	0.7
Teat 2	9.1	9.2	0.0
Teat 3	9.0	7.5	0.3
Teat 4	5.4	8.1	0.0

Table 5.4: Errors between microscribe measurements and triangulations results with offsetting

5.1.3- Summary and discussion:

The results of the first test show an average error of 0.5 mm. The maximum error was 2.8 mm. Such results show that the triangulation routine gives very accurate measurements when the inputs are selected very accurately;

For the second test, the average error was 5.1 mm with a maximum error of 9.2 mm. Such errors are higher than the limit proposed for this project which is of ±5 mm. These high errors can be explained by the limitations of the hardware used. This is also known as the correspondence problem. Figure 5.3 explains how a small deviation when selecting the

corresponding points in the right and left images can cause a higher error in the location of the 3D position of a corresponding point.

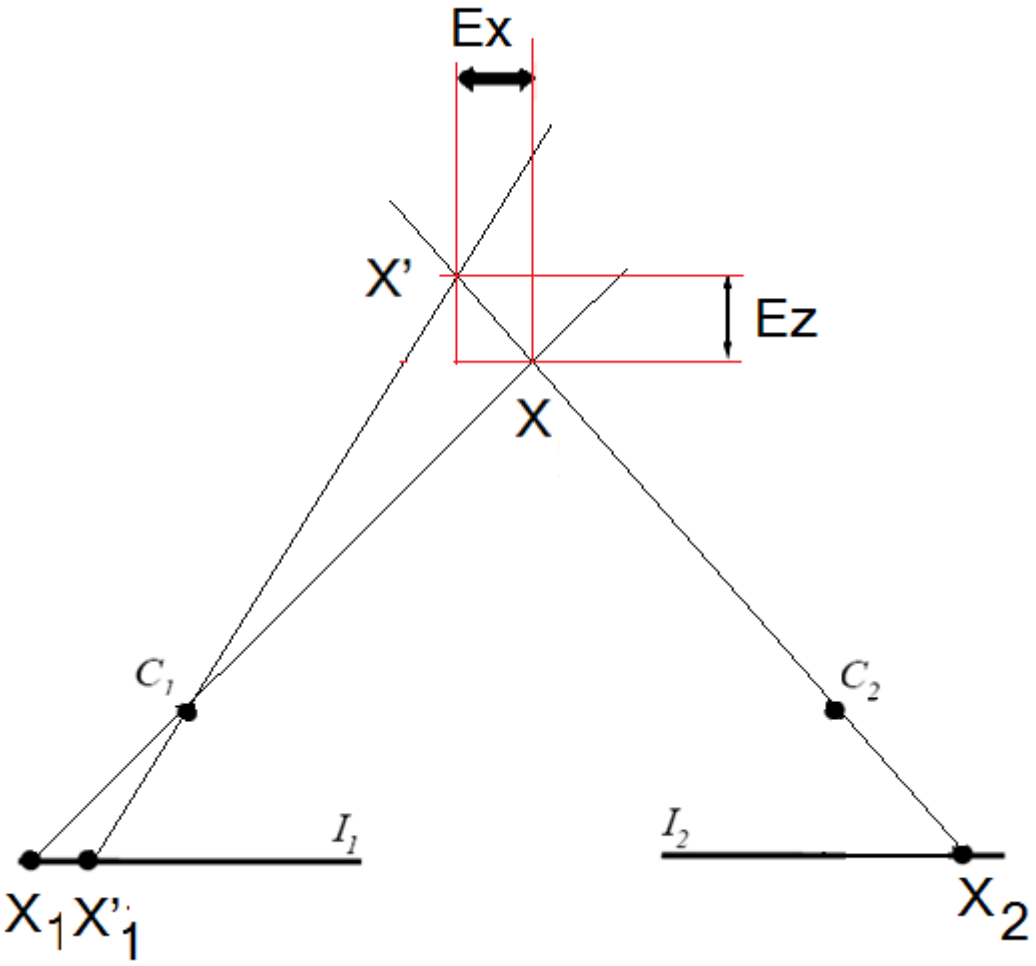


Figure 5.3: Stereo matching error

In Figure 5.3, X'_1 is the offset position of X_1, this small error of detection in the x direction leads to the Ex error in the x direction and Ez error in z direction for the 3D position of the corresponding point X.

Hardware improvement can be introduced to the geometry of the imaging system to decrease the error due to stereo matching. With the current system the focal lengths of the optical cameras were not matched to that of the thermal camera. As a result the teats do not fill a sufficiently large portion of the optical images. A longer focal length for the optical cameras will result in improved image resolution and reduced detection error in the 3D location of object in the scene.

Other techniques to solve the correspondence problem are known as disparity estimation algorithms and their development constitutes one of the most active research areas in

computer vision. However, obtaining robust disparity estimation is difficult, especially in areas of homogeneous colour or occlusion (Kelly, 2007).

5.2- Testing the detection routine

In Section 4, image processing algorithms were created to identify the ends of the teats in the thermal and optical images. These were developed using images of teats in an ideal configuration, that is, unobscured, vertically aligned and without the presence of other heat emitting objects. In this section six scenarios are considered to assess the general purpose application of the image processing algorithms.

- Situation 1- Non-vertical teats: four un-obscured and non touching teats, articulated relative to the vertical, neutral background
- Situation 2- overlapping teats: four un-obscured teats, two are touching and two others are very close to each other, neutral background.
- Situation 3- extra spurious teat shaped object: four clear un-obscured and non touching teats, presence of an object with teat shape in the background
- Situation 4- presence of large heat emitting body at base of teats: four non touching teats, two of them are obscured by an object with same thermal characteristics as the teats; the ends of teats are not covered; only the base. Neutral background
- Situation 5- presence of large heat emitting body at teat ends: four clear un-obscured and non touching teats. Presence of warm object (human arm) in background level with the ends of teats.
- Situation 6- presence of multiple objects at slightly lower temperature: four clear un-obscured and non touching teats. Presence in the background of warm body (human body) with slightly lower thermal characteristics than the teats

When discussing results, a selection of views will be shown, typically showing best and worst case findings using the image processing algorithm.

5.2.1- Situation 1- Non-vertical teats:

Observation of real udders have shown that teat deviation from the vertical is common but usually slight, therefore, for this test, the teats have been articulated with an angle of less than 20 degrees from the vertical. Figure 5.4 shows the identification results for the ends

of the teats on the thermal image. This output is the result of the application of the first step of the main algorithm.

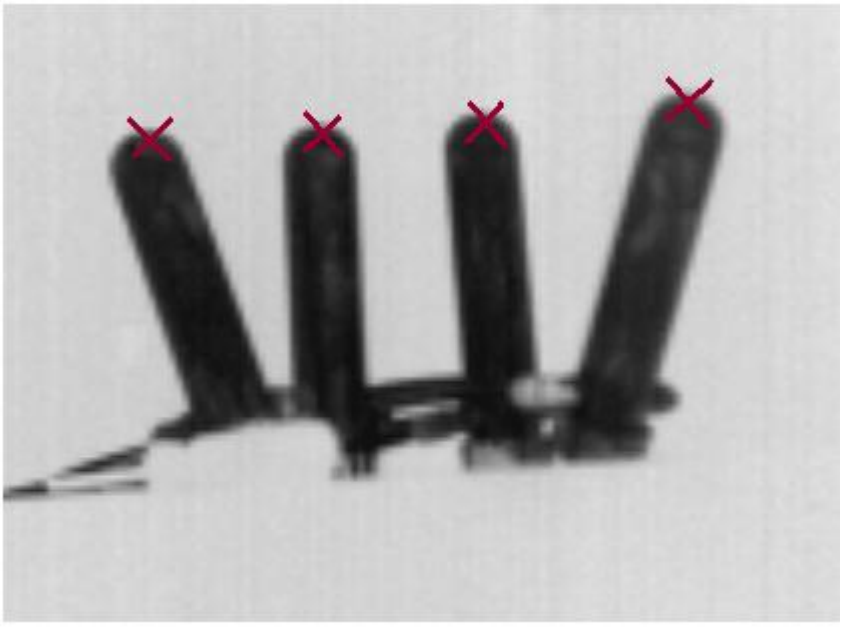


Figure 5.4: Teats detection results in the thermal image- situation 1

As the image shows, and despite the deviation of some of the teats from the vertical, it was possible to successfully identify the ends of the teats. The results of the routine of the second and third steps of the main algorithm are as follows:

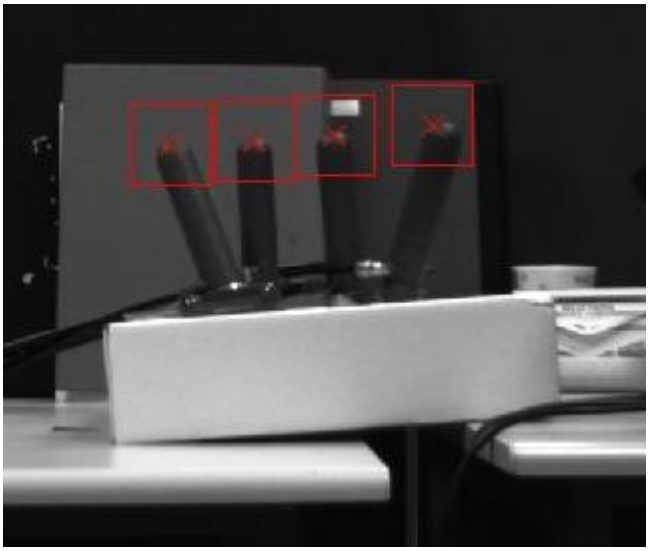


Figure 5.5: Results of homography in left image & ROI for each teat- situation 1

Figure 5.6: Results of homography in right image & ROI for each teat- situation 1

Once the region of interest (ROI) for each teat is specified, the third step of the main algorithm applies a routine to locate the position of the end of the teat more accurately within this region of interest. This routine is applied successively to the four teats in the left image and then in the right image; therefore, the routine is applied eight times. In order to proceed with this part of the algorithm a minimum level of contrast between the teats and the background is required in the optical images. The following images are a selection of the results of this step. For situation one, the ends of the teats have been accurately located with the thermal image, therefore, the results of optical image processing do not change the teat end positions.

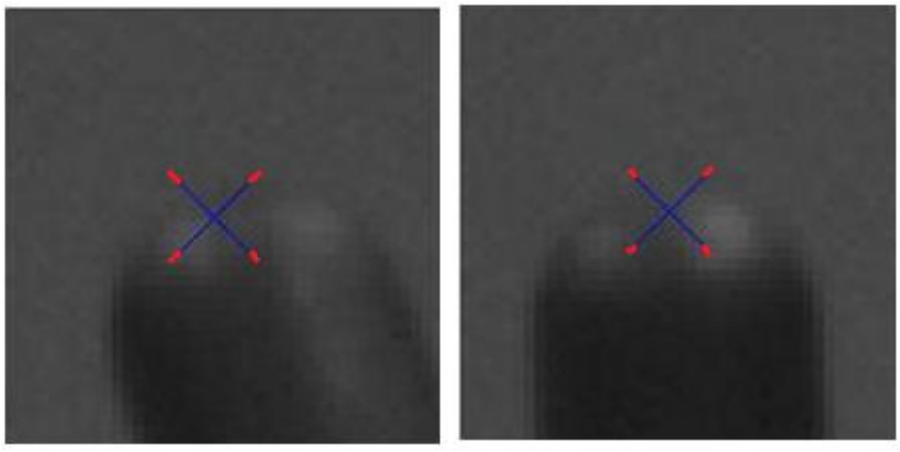


Figure 5.7: Result of end detection of teats 1 & 2 in left image- situation 1

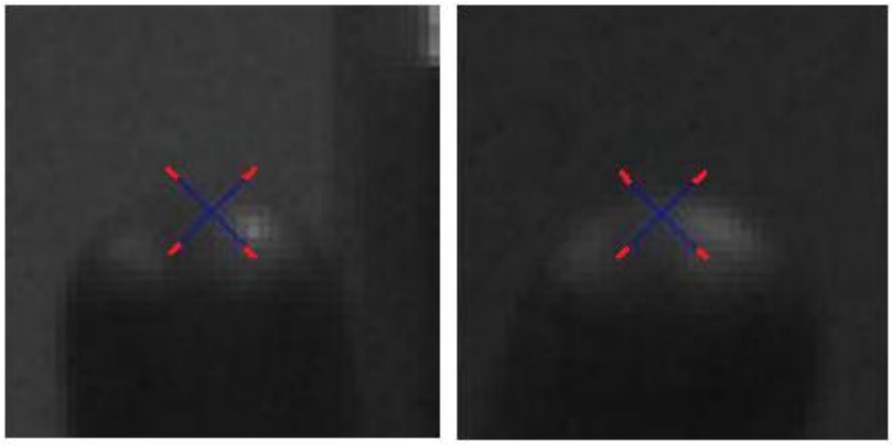


Figure 5.8: Result of end detection of teats 3 & 4 in right image- situation 1

In figures 5.7 and 5.8, the blue crosses are the results of the step four of the main algorithm. The red crosses are the results of the step two, after application of the homographies.

In this situation, situation one, the detection routine successfully identified the end of the four teats. The thermal imaging processing routine identified with high accuracy the locations of the ends of the four teats. The homography matrices successfully identified the corresponding points into the optical images. Therefore, the image processing routine within the optical images successfully identified the ends of teats.

5.2.2- Situation 2- overlapping teats:

In this case the edges of the teats are not clearly defined due to overlapping. Figure 5.9 displays the result of the first step of the process where the ends of the teats are detected in the thermal image.

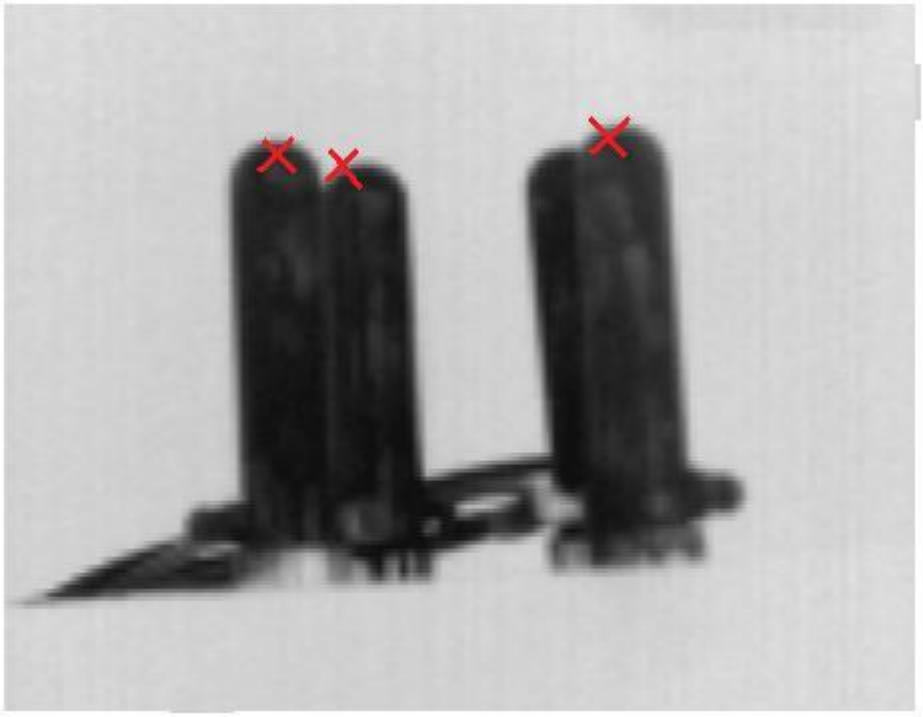


Figure 5.9: Teat detection results in thermal image- Situation 2

As the image shows, the routine identified and located successfully three out of four teats. The third teat from left is partially occluded by the fourth teat and could not be identified. The main algorithm can be improved to overcome such situations by introducing exceptional cases where the main algorithm deviates from the main steps and uses a special algorithm for such configurations. More details about proposed solutions are discussed in section 6.2.

The results of the routine of the second and third steps of the main algorithm are as follows:

Figure 5.10: Results of homography in left image & ROI for each teat- Situation 2

Figure 5.11: Results of homography in right image & ROI for each teat- Situation 2

The next step is the application of the image processing routine within a region of interest for each teat to find more accurately the end of each teat. Figures 5.12 and 5.13 are the results of this step for selected teats in the right and left images.

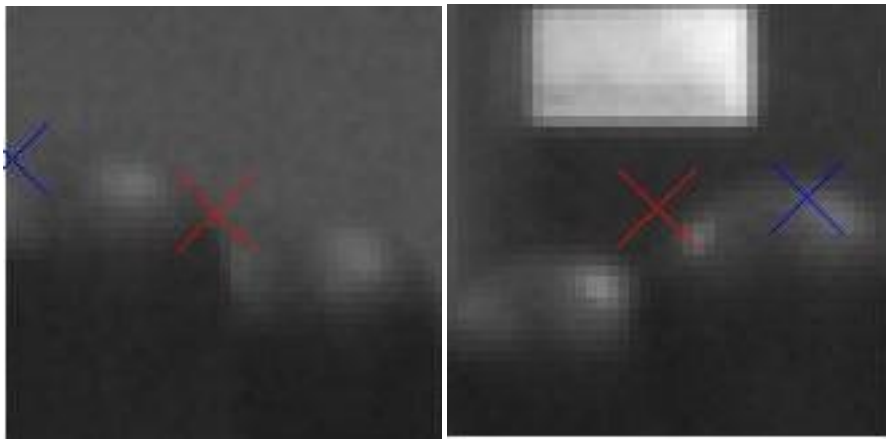


Figure 5.12: Result of end detection of teats 2 & 3 in left image- situation 2

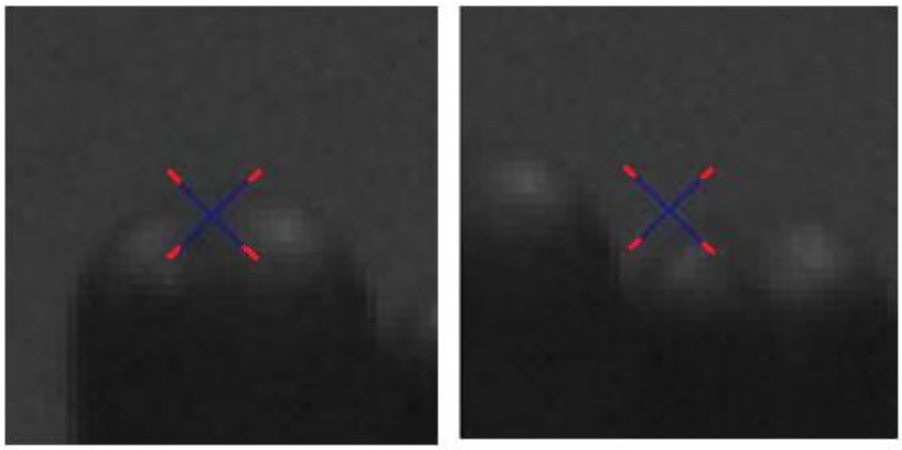


Figure 5.13: Result of end detection of teats 1 & 2 in right image- situation 2

When the final stage of the algorithm was applied to the three identified teats from the thermal image it was found in the right optical image that it was not possible to optimise

the teat positions further. This was due to lack of contrast between the teats and the background and hence the teat locations in the right image are derived from just the homography applied to the thermal image, as Figure 5.13 shows.

However, when this stage was applied to the corresponding teats in the left image, the detection results were improved. In Figure 5.12, the blue crosses are the results of the image processing routine applied to the region of interest and the red crosses are the results of the second step of the main algorithm, that is, the homography from the thermal image to the left image. From this Figure, it is clear that this step has optimised the teat positions further. For example, for the third teat from the left, teat 3, the initial result, the red cross, identified the position of the end of the teat as a point located just out of the side edge of the teat. Step four of the main algorithm rectified this error and located the end of the teat as a point of the teat close to the end edge of the teat, the blue cross. The detection routine applied to the optical image has successfully improved the located positions to varying degrees of the ends of the three initially identified teats.

5. 2.3- Situation 3 – extra spurious teat shaped object:

In this situation, an extra spurious teat shaped object has been placed in the scene (see Figure 5.14) that is dimensionally almost identical to the four phantom teats but is not heated.

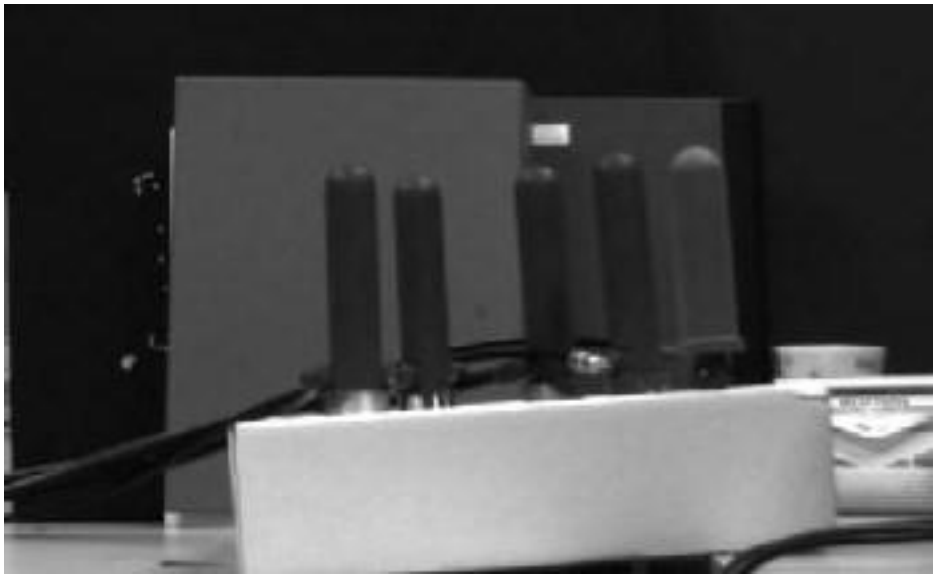


Figure 5.14: Teats configuration with an extra spurious teat shaped object in the far right side

In Figure 5.15, it is apparent that the spurious teat is not detected in the thermal image because it is at background temperature. This vindicates use of a thermal imaging system to reliably identify teats from background detail.

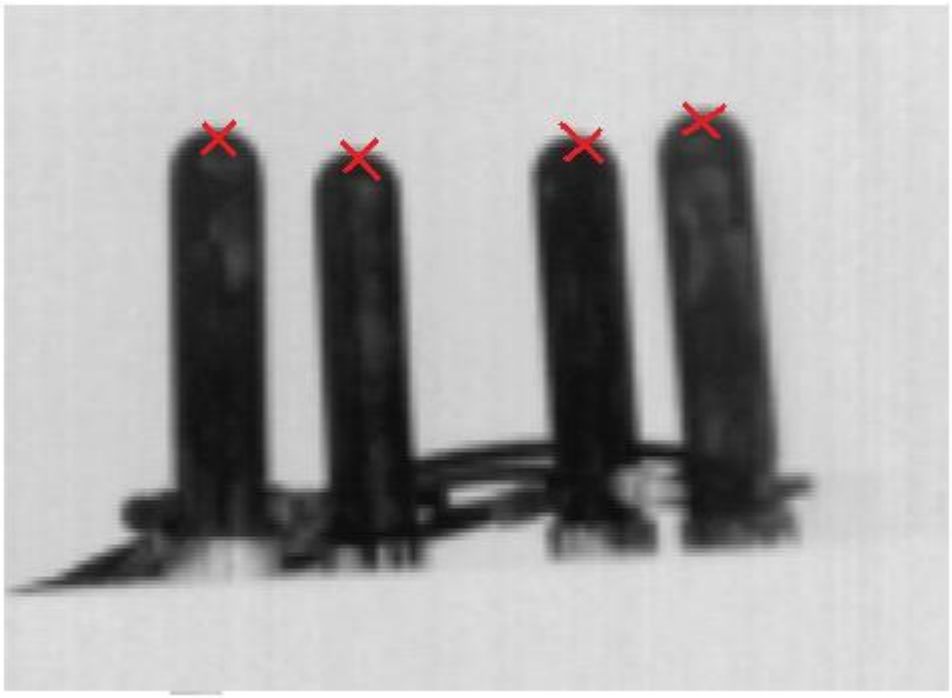


Figure: 5.15: Teat identification, thermal Image- situation 3

The results of the routine of the second and third steps of the main algorithm are as follows:

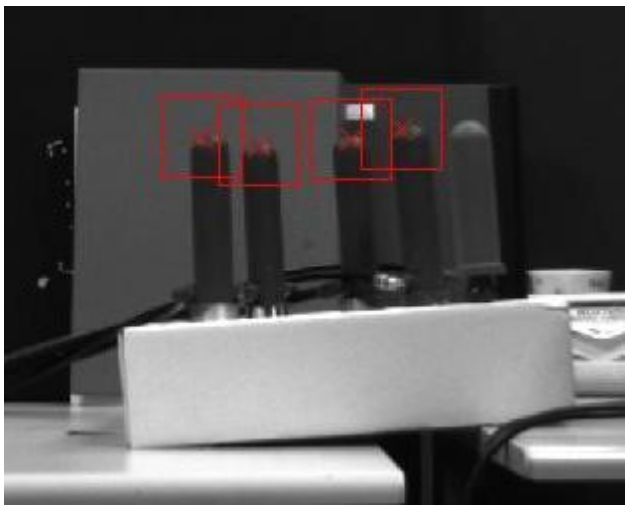


Figure 5.16: Results of homography in left image & ROI for each teat- Situation 3

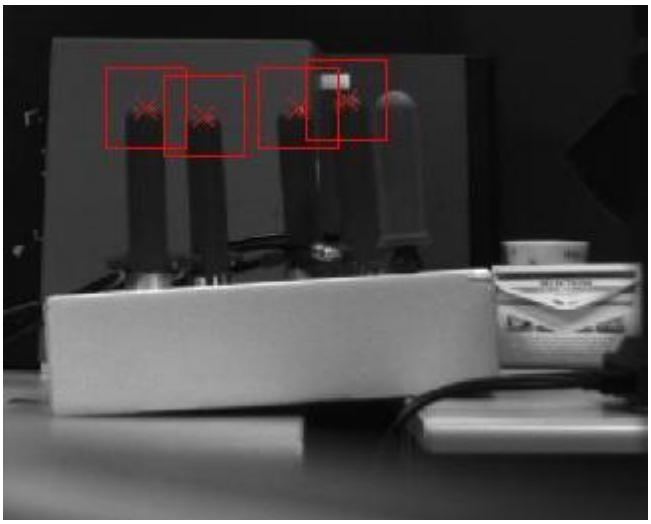


Figure 5.17: Results of homography in right image & ROI for each teat- Situation 3

The next step is the application of the image processing routine within a region of interest for each teat to find more accurately the end of each teat.

It was not possible to further refine teat location using the optical images for two out of the eight teat positions (4 lefts, 4 rights) due to lack of contrast. In these cases the thermally derived points were used.

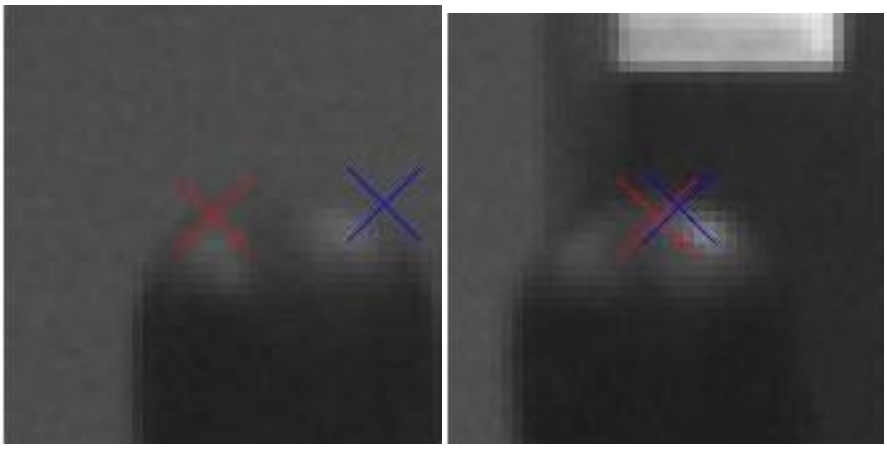


Figure 5.18: Result of end detection of teats 1 & 3 in left image- situation 3

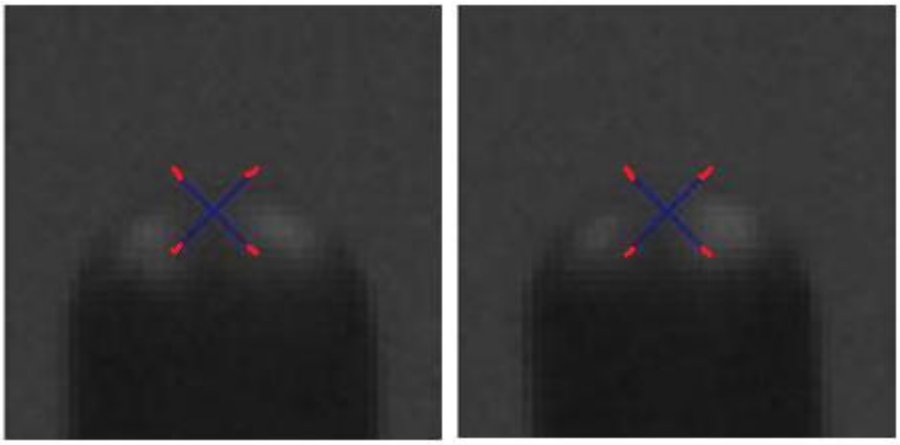


Figure 5.19: Result of end detection of teats 1 & 2 in right image- situation 3

Figure 5.18 shows a sample of the results of the detection routine applied to the regions of interest of the teats in the left image. For the third teat from left, teat 3, the detection routine has slightly improved the accuracy of detection of the end of the teat, the blue cross represent the result of this routine. However, for the first teat from the left, teat 1, the detection routine did not improve on the location as identified using the thermal camera. This is due to the reflection of light on the rounded end of the teat. This depends on the lighting conditions of the scene as well as the reflection characteristics of the experimental teats. Cows' teats will have different surface properties and thus field testing would be required to evaluate this reflection characteristic in an appropriate manner.

Figure 5.19 shows the results of this step with the first and second teats from the left in the right image. The ends of the teats in this image have been already detected with good accuracy after the application of the homography from the thermal image. As a result, the fourth step of the main algorithm did not introduce further processing to the identification algorithm.

5.2.4- Situation 4 –presence of large heat emitting body at base of teats:

In this situation, the system will be assessed with a hot cup situated in the middle of the teats to represent the presence of the udder. The udder is the only other part of the body of the cow that has the same temperature as the teats and therefore it is a source of confusion for the thermal imaging. Figure 5.20 displays the result of the application of the first step of the main algorithm.

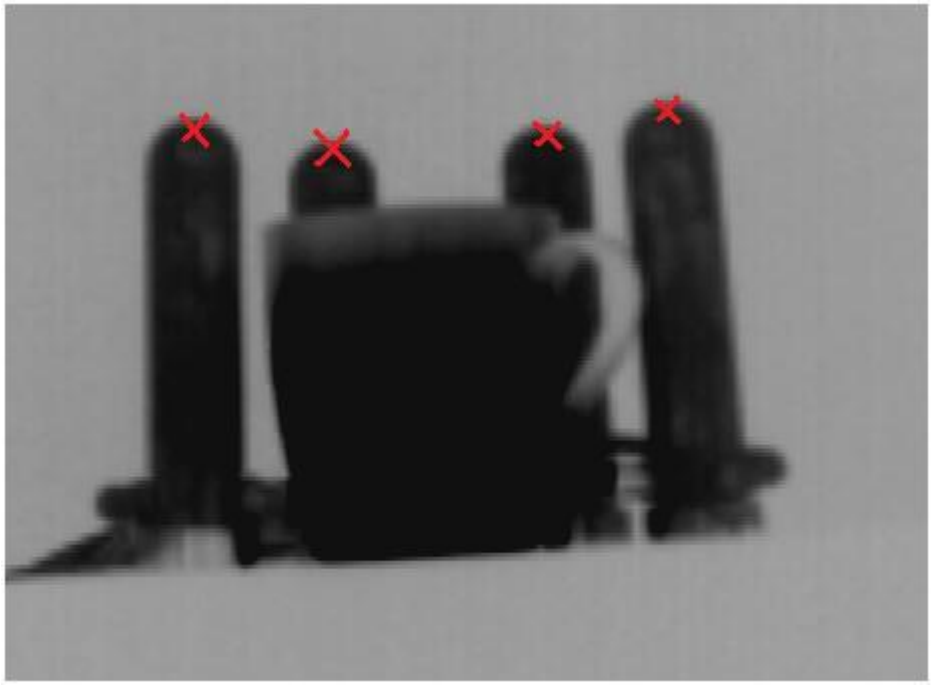


Figure 5.20: Teat detection results in thermal image- Situation 4

As the image shows, the identification and location of the end of the teat was successful. The algorithm identified clearly the ends of the four teats.

The results of the routine of the second step of the main algorithm are as follows:

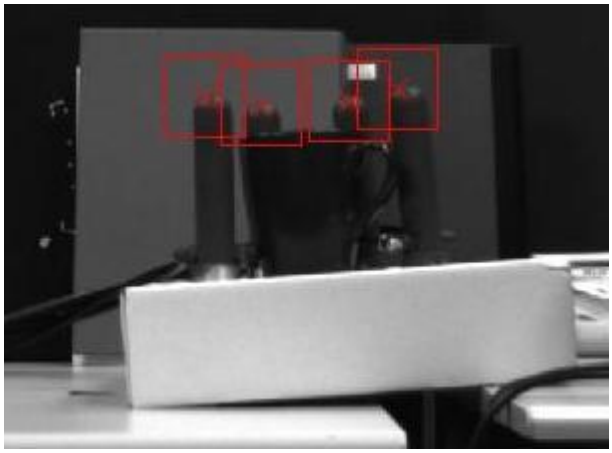


Figure 5.21: Results of homography in left image & ROI for each teat- situation 4

Figure 5.22: Results of homography in right image & ROI for each teat- situation 4

The next step is the application of the image processing routine within a region of interest for each teat to find more accurately the end of each teat.

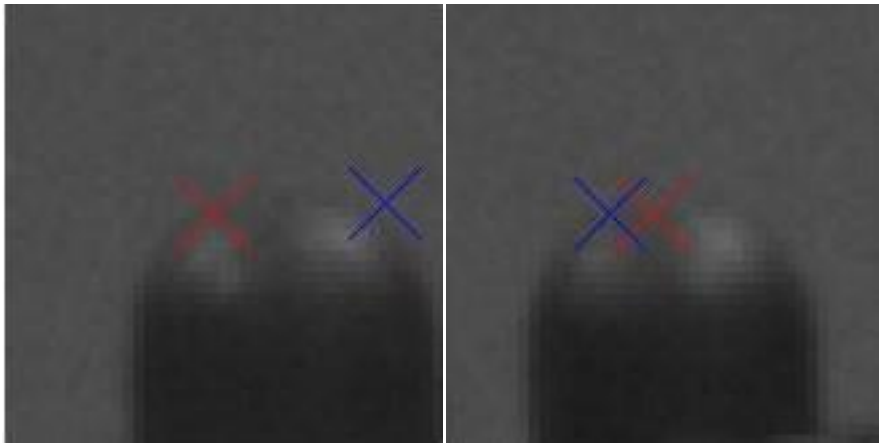


Figure 5.23: Result of end detection of teats 1 & 2 in left image- situation 4

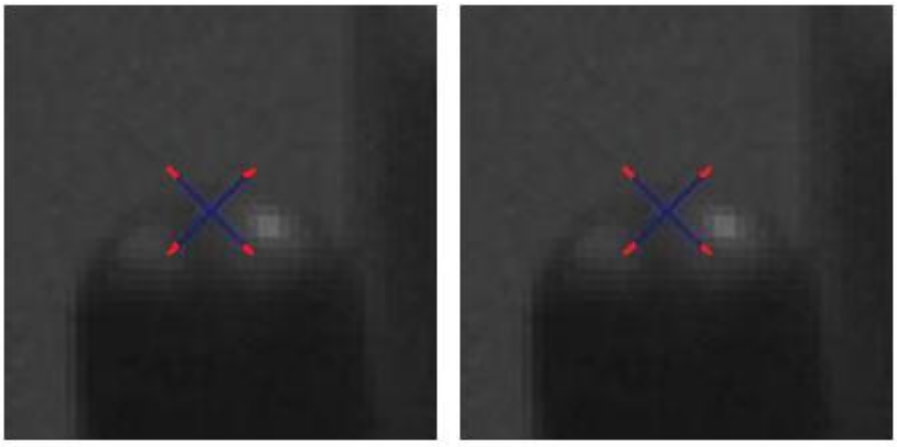


Figure 5.24: Result of end detection of teats 3 & 4 in right image- situation 4

Figures 5.23 and 5.24 display the results of the refinement procedure for selected teats in the right and left images. The results for this situation are similar to the results obtained in situation 3. The only difference is the introduction of a cup positioned in between the teats, however, the cup does not interfere with the view of the optical cameras and therefore the results of the fourth step of the main algorithm are the same as situation 3. In this situation, and except for the case of the first teat in the left image which is affected by light reflection, the routine has improved the accuracy of detection of the ends of the teats where required.

5.2.5- Situation 5- presence of large heat emitting body at teat ends:

The results of the previous situation prompted the scenario for situation 5. Here, as opposed to the preceding case, the ends of the teats are not clearly identifiable in the thermal image. This is due to the presence of a warm body (human arm) in the background near the end of the teats, see Figure 5.25.

Figure 5.25: Heat emitting body near teat ends

The following image shows the result of the application of the first step of the main algorithm to this situation.

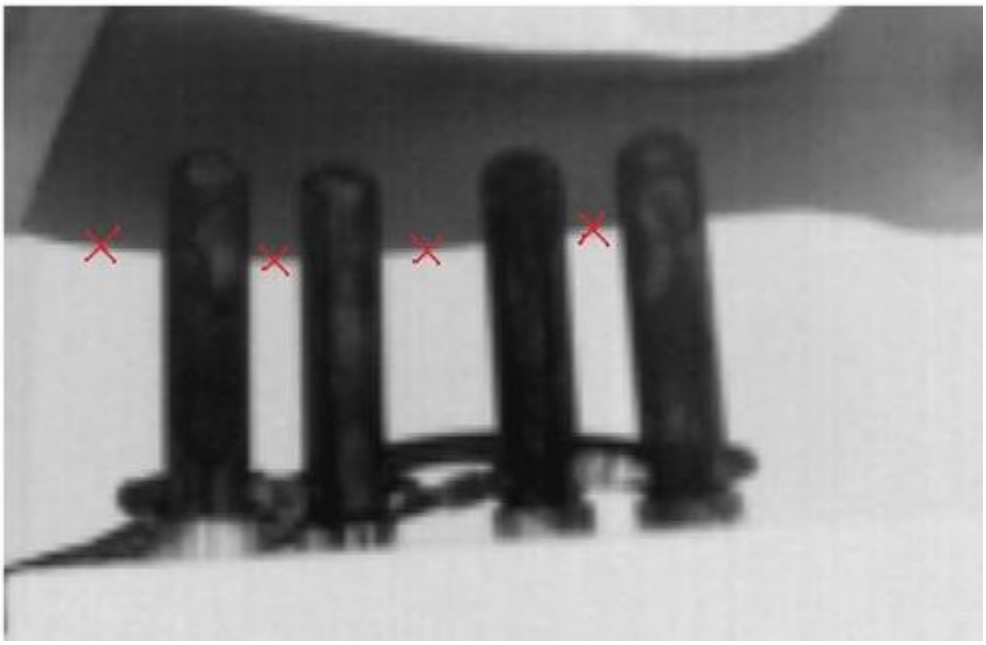


Figure 5.26: Teat identification thermal image- situation 6

The routine did not identify correctly any of the four teats. The system has misidentified the ends of the teats with the border of the human arm. This indicates that the background in the vicinity of the teat ends must emit less heat radiation than the teats for reliable identification. The second step, homography from thermal image to optical images, identifies points of the border of the human arm as the ends of the teats (see Figure 5.27) and therefore, any further steps are inappropriate.

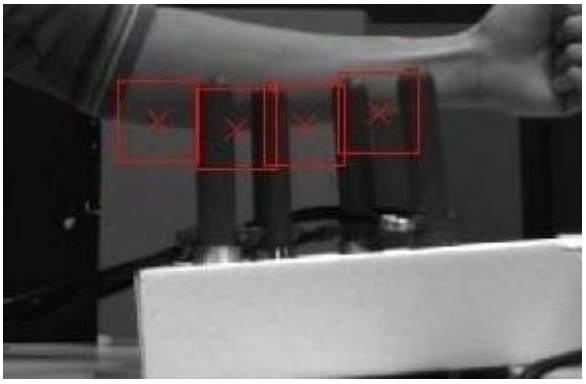


Figure 5.27: Results of step 3 and 4 in left image- Situation 5

5.2.6- Situation 6- presence of multiple objects at slightly lower temperature:

This situation reflects more closely the real scenario with the presence of other parts of the animal such as the torso and the legs which are somewhat cooler than the udder and teats but above the background temperature. Figure 5.28 displays the result from the first step in the main algorithm.

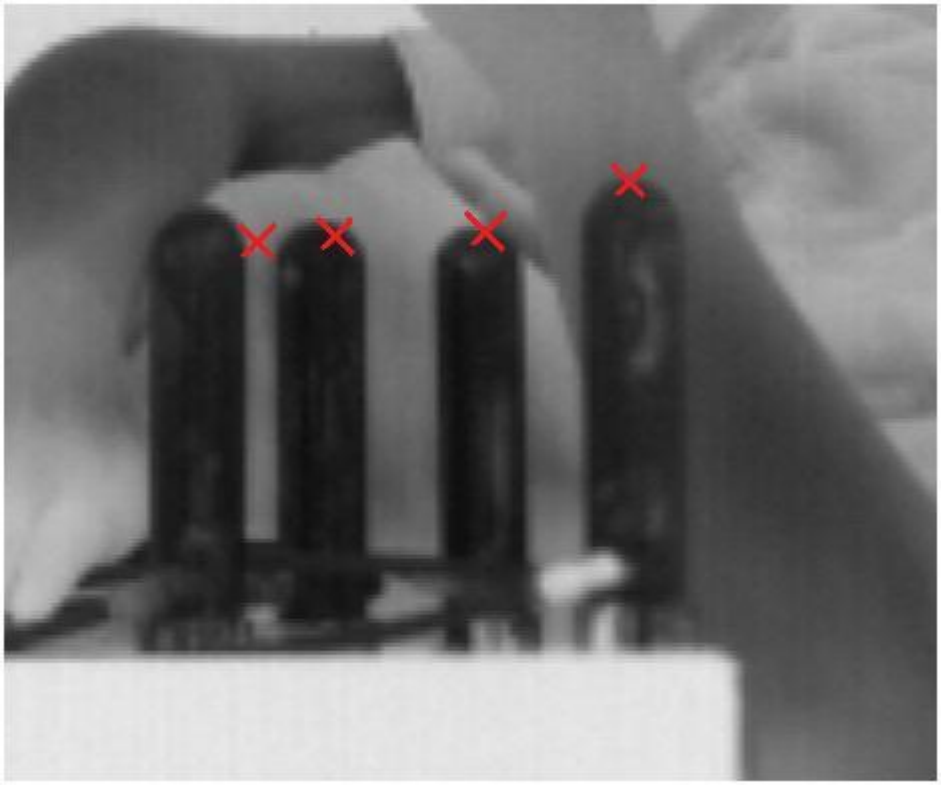


Figure 5.28: Teat identification thermal image- situation 6

As the image shows, the identification process encountered some problems, only identifying three teats correctly. The results of the routine of the second step of the main algorithm are as follows.

Figure 5.29: Results of homography in left image & ROI for each teat- Situation 6

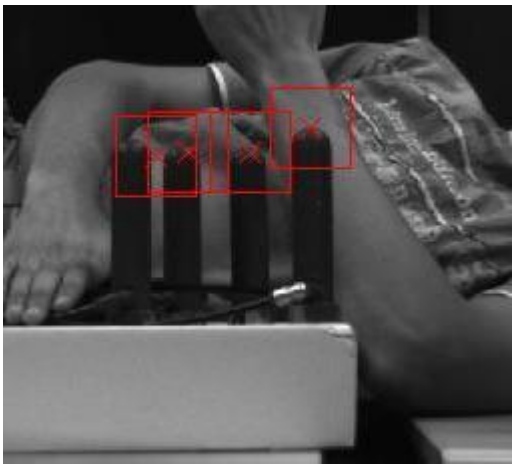


Figure 5.30: Results of homography in right image & ROI for each teat- Situation 6

The next step is the application of the image processing routine within a region of interest for each teat to find more accurately the end of each teat.

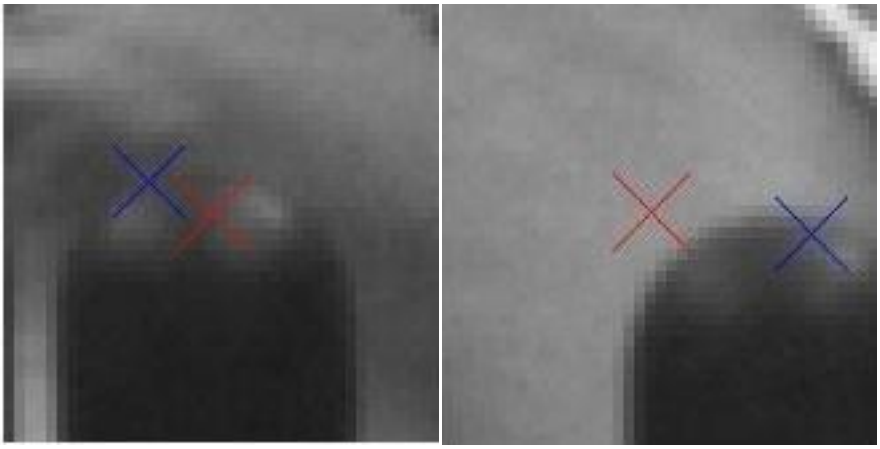


Figure 5.31: Result of end detection of teats 2 & 4 in left image- situation 6

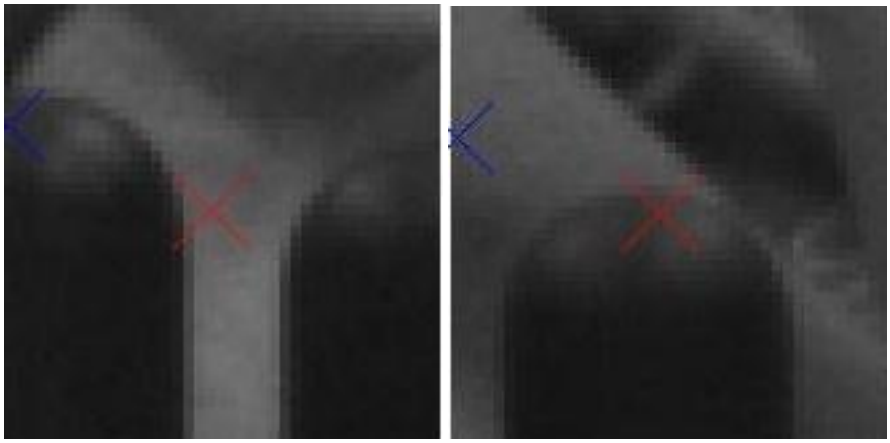


Figure 5.32: Result of end detection of teats 1 & 3 in right image- situation 6

Figures 5.31 and 5.32 display a selection of the results obtained during the refinement process. In Figure 5.31, the detection routine has corrected the location of teat 4 (blue

cross) in the left image. From the same image, the detection routine did not improve the location of teat 2; the new location, the blue cross, is closer to the edge of the teat than derived with the thermal camera (red cross). Again this phenomenon is related to the light reflection characteristics of the teats in the optical images.

In Figure 5.32, the refinement process has improved the estimate of teat 1 but increased the error for teat three. This shows that while the thermal camera improves the system's ability to identify the teats, good optical contrast is still required around the teat end locations for the refinement process to function adequately. Software and hardware solutions to improve these issues are discussed in the section 6.2.

5.2.7- Summary and discussion:

In this section the image processing algorithms created and detailed in chapter 4 were tested. Six scenarios were considered. In the first situation, the teats have been slightly deviated from the vertical. The results have shown that the detection routine successfully identified the end of the four teats. The accuracy of the detection was very high and the detection error could be considered as nil (considering that detection of any point of the end of the teat is a perfect result).

In the second situation, the teats are overlapping; two teats are touching and the fourth teat is partially occluding the third teat. The routine successfully detected three teats out of four; the occluded teat was missed. The image processing routine applied to the thermal image has detected the three teats and the detection routine applied to the optical image successfully improved the located positions to varying degrees of the ends of the three initially identified teats. To overcome the issue of the occluded teat, the main algorithm can be improved by introducing exceptional cases where the main algorithm deviates from the main steps and uses particular steps for such configurations. Such particular programming step would rely on the data collected by the optical cameras and a historical data associated with machine learning algorithms: once the system is installed with an automated milking system, the positions of the teats of each cow will be collected and saved in a data base. Observations with cows have shown that positions of teats of a cow vary slightly between two successive milkings. This data base will be updated every time the cow enters the milking system and the Machine learning algorithms will evaluate the positions of the teats depending on inputs such as last recorded position and cow's lactation period.

The third situation vindicated the use of thermal imaging; an extra spurious teat shaped object was introduced in the scene, however, the thermal camera did not detect this object. Such an object would be a source of confusion for the optical cameras and can be identified as a teat. The results of the detection routine for this situation showed that the four teats were detected with varying accuracy. The image processing routine applied to the optical images did not improve the location of the end of the teat for one case. This is due to the reflection of light on the rounded end of the teat. This depends on the lighting conditions of the scene as well as on the reflection characteristics of the experimental teats. Cows' teats have different surface properties and thus field testing would be required to evaluate this reflection characteristic in an appropriate manner.

In situation 4, a hot cup was placed in the middle of the teats to represent the presence of the udder. The udder is the only other part of the body of the cow that has the same temperature as the teats and therefore it is a source of confusion for the thermal imaging. The results of this situation have shown that the detection routine just needs an unobscured view of the end of the teat in the thermal camera to detect its position. The udder (hot cup in this case), does not interfere with the view of the end of teats. Therefore, the ends of the four teats were successfully detected. The accuracy of the detection varies from a teat to another.

The results of the fourth situation prompted the scenario for situation 5. A warm body (human arm) was placed in the background of the end of the teats. As opposed to the preceding case, the ends of the teats are not clearly identifiable in the thermal image and the routine has completely misidentified the four teats. This is confirming that in order to reliably detect the four teats, the thermal camera needs a clear view of the ends of the four teats.

The sixth and last situation reflects more closely the real scenario with the presence of other parts of the animal such as the torso and the legs which are somewhat cooler than the udder and teats but above the background temperature. Multiple objects at slightly lower temperature than the teats have been placed in the background. By and large the optical processing improved the detection of teat location although for one teat the optical processing increased the position error. This shows that while the thermal camera improves the system's ability to identify the teats, good optical contrast is still required around the teat end locations for the refinement process to function adequately. Data fusion and mutual thresholding (O'Conaire, 2007) are Software solutions that could be considered to improve these issues.

CHAPTER 6

Conclusions and Future Work

6.1- Conclusions:

The objective of this work was to investigate the use of a hybrid sensing system for teat detection in automatic milking. Previous work had shown that a purely optical system was prone to error arising out of an inability to correctly identify each teat. The strong heat signature from the udder region prompted the addition of a thermal imaging system for use in conjunction with the optical system. The overall system was evaluated from the point of view of accuracy and robustness. The higher pixel density of the optical system was required to give the appropriate positional accuracy while the thermal system was used to create a small search region when identifying the teat ends in the optical images. From system testing the following conclusions can be drawn.

If the correct pixel correspondences for the teat ends are identified, the accuracy available from the optical system is within the required level of ±5mm. In the tests conducted here this could have been improved by using optical cameras with the appropriate focal length so that the teat region occupied a larger portion of the image.

Small disparities of pixel correspondences can lead to large errors. For the hardware set-up used in the trials, one pixel error led to 9mm inaccuracy. This highlights the need for robust teat identification.

Using the thermal imaging system, it was possible to identify teats even when spurious teat shaped objects were present in the scene. This validates the central objective of this work.

During the positional determination stage of the process using the optical cameras, the following conditions are necessary

i) An unobstructed view of the teats

ii) Adequate lighting

iii) Contrast between the teats and the background in the teat region.

In the case where a teat is obstructed it may be possible to use historical data on the relative positions of teats to predict the unknown position from the information obtained from the visible teats.

6.2- Future Work:

For future work, it is proposed to develop the system using live animal trials. That is, to set up the hardware in a milking parlour and evaluate the system under more realistic conditions. From a hardware point of view, the following should be considered

a) The focal length of the optical and thermal systems should be matched so that the teats occupy the appropriate region in all images.
b) Adequate lighting should be provided so that small apertures can be used with the optical cameras, this will ensure sharper images are obtained which will improve pixel correspondence between images
c) Some adaptation to the milking stall will be required so that good contrast is obtained between the teats and the background.

As regards the software, the algorithms developed in this work were based around the identification of a thermal phantom udder in the laboratory. Some of the techniques may prove useful in future work but it will only be possible to ascertain the degree of further development required once testing begins in situ with live animals in the milking parlour.

REFERENCES:

Agriculture in Ireland ,2002, *Teagasc Web site*, [Online], Available: http://www.teagasc.ie/agrifood/ [23 March 2009].

Akermann, D. ,1979, 'Verfahren und Vorrichtung zum Melken. (Procedure and device for Milking)', *Deutsche Offenlegungsschrift 28 (49)*, p. 227.

Andersson, L. and Nilsson, M. ,1998, *Apparatus and method for recognising and determining the position of a part of an animal*, WO/1998/047348.

Artmann, R. ,1997, 'Sensor systems for milking robots', *Computers and Electronics in Agriculture*, pp. 19-40.

Artmann, R. and Schillingmann, D. ,1989, 'Automation of milking by using robots and electronics', Proceedings of second International Conference, Agrotique, Bordeaux, 330-348.

Artmann, R. and Schillingmann, D. ,1990, 'Entwicklungsstand von Melkrobotern.(State of the art of milking robots)', *Landtechnik 45*, p. 437–440.

Artmann, R., Schillingmann, D., Krdger, G. and Bollmann, B. ,1990, 'Lokalisierung der Zitzenposition mittels Ultraschall und Bildverarbeitungssystem', Robotereinsatz in der Landwirtschaft am Beispiel des Melkens, Vol. 9, VDI/MEG Kolloquium, Dusseldorf, 126-152.

Ben Azouz, A., Hunt Duffy, A., Corcoran, B., Esmonde, H. and O'Callaghan, E. ,2008, 'Development of a Teat Sensing System for Robotic Milking', In Proceedings of The 11th Mechatronics Forum Biennial International Conference (MECH 2008), Limerick, Ireland.

Bergerot, P., Baylou, P. and Ordolff, D. ,1989, 'Systeme de pose automatique des gobelets trayeurs', Agrotique 89, Bordeaux, 317-330.

Bouguet, J.Y. ,2008a, *First calibration example*, 2June, [Online], Available: http://www.vision.caltech.edu/bouguetj/calib_doc/htmls/example.html [23 March 2009].

Bouguet, J. ,2008b, *Home Page*, 2June, [Online], Available: http://www.vision.caltech.edu/bouguetj/calib_doc/index.html [23 March 2009].

Burnay, S., Williams, T.L. and Jones, C.H.N. ,1988, 'Applications of thermal imaging', *Optics and Lasers in Engineering*, pp. 145-146.

DeLaval ,2008, *DeLaval VMS in detail*, 27June, [Online], Available: http://www.delaval.com/Products/Automatic-Milking-Robotic-milking/DeLaval-VMS-in-detail/default.htm?wbc_purpose=BasicA [23 March 2009].

DeLaval herringbone ,2007, *DeLaval EnDurance herringbone*, 26January, [Online], Available: http://www.delaval.co.uk/NR/rdonlyres/F302EEE9-730C-46FB-A843-ABA4E68F0E24/0/EnduranceHD.jpg [23 March 2009].

DeLaval Rotary ,2008, *DeLaval Rotary*, 26June, [Online], Available: http://www.delaval.com/NR/rdonlyres/4295E840-4045-4127-9F47-517EAAC40904/0/Rotary_PER_web.jpg [23 March 2009].

DeLaval VMS ,2008, *DeLaval VMS Photo*, 27June, [Online], Available: http://www.delaval.com/NR/rdonlyres/7862CE85-5E1A-45BD-BDD9-060C8F98BAD8/0/DeLavalVMSFront.jpg [23 Mar 2009].

Dispatch&Garlick ,2009, *Magnum Rotary cow sheds*, [Online], Available: http://www.dispatchgarlick.co.nz/Images/mexico1.jpg [23 March 2009].

Duck, M. ,1992, 'Evolution of Duvelsdorf milking robot', *Prospects for Automatic Milking* , pp. 49-54.

Duffy, A.H. ,2006, *Teat Detection for an Automated Milking System*, MEng Thesis: Dublin City University.

Faugeras, O. ,1993, *Three Dimensional Computer Vision A Geometric Viewpoint*, Massachusetts: Massachusetts Institute of Technology.

FLIR ,2004, *A20 M Operator's Manual*, FLIR Systems.

Forsythe, D. and Ponce, J. ,2003, *Computer Vision - A Modern Approach*, Upper Saddle River, New jersey: Prentice Hall.

Fu, K., Gonzales, R. and Lee, C. ,1987, *Robotics: Control, Sensoring, Vision and Intelligences*, New York: McGraw-Hill.

Fullwood ,2007, *Merlin Automated Milking System*, [Online], Available: http://www.fullwood.com/milking_products/139/ [23 March 2009].

Gabler, E. ,1971, *Melkeinrichtung, vorzugsweise furgroße Milchviehbestande, (Milking device, preferably for large herds)*, DDR patent 82 592.

Gonzalez, R.C., Woods, R.E. and Eddins, S.L. ,2004, *Digital Image Processing using Matlab*, New Jersey: Pearson Prentice Hall.

Hartley, R. and Zisserman, A. ,2003a, '3D Reconstruction of Cameras and Structure', in *Multiple View Geometry in Computer Vision*, 2nd edition, Cambridge University Press.

Hartley, R. and Zisserman, A. ,2003b, 'Camera Models', in *Multiple View Geometry in Computer Vision*, 2nd edition, Cambridge: Cambridge University Press.

Hartley, R. and Zisserman, A. ,2003c, 'Computation of the Camera Matrix P', in *Multiple View Geometry in Computer Vision*, 2nd edition, Cambridge University Press.

Hartley, R. and Zisserman, A. ,2003d, 'Epipolar Geometry and The fundamental Matrix', in *Multiple View Geometry in Computer Vision*, 2nd edition, Cambridge University Press.

Hartley, R. and Zisserman, A. ,2003e, 'Estimation- 2D Projective Transformation', in *Multiple View Geometry in Computer Vision*, 2nd edition, Cambridge: Cambridge University Press.

Immersion ,2009, *Immersion Corporation Web site*, [Online], Available: http://www.immersion.com/ [23 March 2009].

Internal Report ,2006, *Existing Technologies and Patent Reviews of AMS*, Department of Mechanical and Manufacturing Engineering, Dublin City University.

Jago, J., Ohnstad, I. and Reinemann, D.J. ,2007, 'Labor Practices and Technology Adoption on New Zealand Dairy Farms', Sixth International ASABE Dairy Housing Conference, Minneapolis, Minnesota.

Jensen, T. ,2004, 'Expectations of automatic milking and the realized socio-economic effects', Automatic milking, a better understanding, Lelystad, 78-79.

Jones, B.F. and Plassmann, P. ,2002, 'Digital infrared thermal imaging of human skin', *IEEE engineering in medicine and biology magazine*, Nov/Dec, pp. 41-18.

K. Koning, J.R. ,2004, 'Automatic milking: State of the art in Europe and North America', Automatic milking, a better understanding, Lelystad, 27-37.

Kelly, P. ,2007, *Pedestrian Detection and Tracking using Sterep Vision Techniques*, PhD Thesis: Dublin City University.

KwaZulu ,1999, *Dairying in KwaZulu-Natal*, [Online], Available: http://agriculture.kzntl.gov.za/publications/production_guidelines/dairying_in_natal/dairy7_1.htm [23 March 2009].

Lely ,2008, *Lely Milking Equipment*, [Online], Available: http://www.lely.com/en/dairy-equipment/milking-equipment/index.jsp [23 March 2009].

Mathijs, E. ,2004, 'Socio-economic aspect of automatic milking', Automatic milking, a better understanding, Lelystad, 46-55.

MathWork Dilate ,2009, *Imdilate: Dilate image*, [Online], Available: http://www.mathworks.com/access/helpdesk/help/toolbox/images/index.html?/access/helpdesk/help/toolbox/images/f3-23960.html&http://www.mathworks.com/products/image/ [23 March 2009].

MathWork Filter ,2009, *ordfilt2: 2-D order-statistic filtering*, [Online], Available: http://www.mathworks.com/access/helpdesk/help/toolbox/images/index.html?/access/helpdesk/help/toolbox/images/f3-23960.html&http://www.mathworks.com/products/image/ [23 March 2009].

Montalescot, J. ,1987, 'Robotisation de la traite: Des Travaux déjà bien avancés.(Robotic milking: research already has made good progress)', *Revue laitière élevage 30 (3)*, pp. 101-103.

Moorepark ,2009, *Teagasc Moorepark Food Research Centre*, [Online], Available: http://www.teagasc.ie/research/MFRC.asp [23 March 2009].

Mottram, T. ,1997, 'Requirements for teat inspection and cleaning in automatic milking systems', *Computers and electronics in agriculture, Issue 17*, p. 63 – 77.

Navitar ,2009, *Navitar Video Lenses*, [Online], Available: http://machinevision.navitar.com/catalog/product/?id=182&c=180 [23 March 2009].

Newman, E.A. and Hartline, P.H. ,1982, 'The infrared “vision" of snakes', *Scientific American*, 10 March, pp. 116-127.

Notsuki, I. and Ueno, K. ,1977, *System for Managing Milking Cows in Stanchion Stool*, Patent 4 010 714.

NZDB ,1984, *Artificial Breeding Questionnaire*, New Zealand Dairy Board.

O'Brien, B., O'Donovan, K., Gleeson, D., Kinsella, J. and Ruane, D. ,2002, 'Profiling the working year on Irish dairy farms - identification of some work areas towards

improvement in efficiency', Proceedings of Irish Grassland Association Dairy Conference, Cork, 112-124.

O'Callaghan, E., O'Brien, B., Gleeson, D. and O'Donovan, K. ,2001, *Milking facilities and labour*, [Online], Available: http://www.teagasc.net/publications/2001/ndc/ndc-callaghan.asp [23 March 2009].

O'Conaire, C. ,2007, *Adaptive Detection and Tracking using Multimodal Information*, PhD Thesis: Dublin City University.

O'Conaire, C., Cooke, E., O'Connor, N., Murphy, N. and Smeaton, A.F. ,2005, 'Fusion of infrared and visible spectrum video for indoor surveillance', International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Montreux, Switzerland.

Ordolff, D. ,1984, 'A system for automatic teat cup attachment', *Eng.Res.*, pp. 65-70.

Ordolff, D. ,2001, 'Introduction of electronics into milking Technology', *Computers and Electronics in Agriculture*, pp. 125-149.

Parker, J.R. ,1997, *Algorithms for image processing and computer vision*, New York: John Wiley & Sons, Inc.

Pilkington ,2009, *Pilkington K Glass™*, [Online], Available: http://www.pilkington.com/applications/products2006/english/bybenefit/thermalinsulation/kglass/default.htm [23 March 2009].

Pluim, J., Maintz, J. and Viergever, M. ,2003, 'Mutual-information-based registration of medical images: a survey', *Medical Imaging, IEEE Transactions 22 (8)*, Aug, pp. 289-1004.

Prosilica ,2005, *EC1350 User Manual*, 19October, [Online], Available: http://www.prosilica.com/products/manuals/700006AB-EC1350%20User%20Manual.pdf [23 Mar 2009].

Review Document ,2006, *Teat Location for a Rotary Carousel Automatic Milking System*, Internal Report, Department of Mechanical and Manufacturing Engineering: Dublin City University.

Scheidemann, B. ,1990, 'Ein kartesisch arbeitender Roboter-Aufbau und Erfahrungen (A Cartesian robot-structure and experiences)', Robotereinsatz in der Landwirtschaft am Beispiel des Melkens H.9, VDI/MEG Kolloquium, Dusseldorf, 221-227.

Sonka, M., Hlavac, V. and Boyle, R. ,1999, *Image Processing, Analysis and Machine Vision*, 2nd edition, PWS Publishing.

Torsius, A. ,1987, *Milking apparatus*, EP0213660A1.

Viola, P.A. ,1995, *Alignment by Maximization of Mutual Information*, Phd thesis: Massachusetts Institute of Technology, Massachusetts (MA), USA.

Wade, K., van Asseldonk, M., Berentsen, P., Ouweltjes, W. and Hogeveen, H. ,2004, 'Economic efficiency of automatic milking systems with specific emphasis on increases in milk production', Automatic milking, a better understanding, Lelystad, 62-67.

Woolford, M., Claycomb, R., Jago, J., Davis, K., Ohnstad, I., Wieliczko, R. and Bright, P.C.a.K. ,2004, 'Automatic dairy farming in New Zealand using extensive grazing systems', Automatic milking, a better understanding, Lelystad, 280-285.

Xu, B. and Aneshansley, D. ,1991, 'A vision-guided automatic dairy cow's teat-cup attachment system', ASAE International Winter Meeting, Chicago, Illinois, No. 913513.

Appendices

A P P E N D I X A

Software Code

A.1- Main algorithm:

The following program is the matlab code developed for the main algorithm of teat identification (section 4). The steps of this algorithm are as follows:

- Image reflection
- Edge detection (thermal image)
- Find verticals (thermal image)
- Find horizontals (thermal image)
- Find max of verticals (thermal image)
- Find middle of horizontals (thermal image)
- Find end of teats (thermal image)
- Apply the homography to the thermal image and locate the position of the teats in the optical images
- Find end of the teats in the left optical image
- Find end of the teats in the right optical image

```
Function
[teat_left,teat_right,hiddenteats,X3D_teats_microscribe]=getteats(thermal,left,right,Hright
,Hleft,TransfMatrix,om,T,fc_left,cc_left,kc_left,alpha_c_left,fc_right,cc_right,kc_right,al
pha_c_right)
% Detect the position of end of the teats in the thermal image
R=reflection(thermal); % reflect image
thermal=R;
BW= edge(thermal,'canny', [],12); % detect edges in the thermal image
BW(:,631)=0;
domainvertical=[0 0 0 0 1 0 0 0 0;0 0 0 0 1 0 0 0 0;0 0 0 0 1 0 0 0 0;0 0 0 0 1 0 0 0 0;0 0 0
0 1 0 0 0 0;0 0 0 0 1 0 0 0 0;0 0 0 0 1 0 0 0 0;0 0 0 0 1 0 0 0 0];
BWv = ordfilt2(BW, 1, domainvertical); % find verticals
domainhoriz=[0 0 0;0 0 0;1 1 1;0 0 0;0 0 0];
BWh=ordfilt2(BW,1,domainhoriz); % find horizontals
```

```
maxvertical=maxverticals(BWv); % Find max of verticals
p=round(mean(maxvertical,1));
maxhoriz=horizontals(BWh,p); %find max of horizontal
[teats,hiddenteats]=endteats(maxhoriz); % Find ends of teats
plotpoints(thermal,teats)

% apply homography to thermal results,get x,y of teats in optical images:
 [teats_left,teats_right]=getteatsOptical(teats,Hright,Hleft);
 teats_left=round(teats_left);
 teats_right=round(teats_right);

% plot teats position on the optical images:
    plotpoints(left,teats_left')
    figure()
    plotpoints(right,teats_right')

% locate each teat in left optical image
teat_left={1,4};
st=size(teats,1);
%if st==4
   for i=1:st
   [imageoptl{i},BWl{i},teat_left{i}]=locatinoptical(left,teats_left,i);
   end
 % locate each teat in right optical image
   teat_right={1,4};
   for i=1:size(teats,1)
   [imageoptr{i},BWr{i},teat_right{i}]=locatinoptical(right,teats_right,i);
   end

    plotpoints(left,teats_left')
    figure()
   plotpoints(right,teats_right')

end
```

A.2- Routines applied to thermal image:

A.2.1- Reflection of the thermal image:

```
%Program for one sided image reflection along a Line
function R = reflection(I)
%I - Image to be Reflected
%L - Line position
%R - Reflected image
%eg : R = reflection(I,128);
% L should be between 1 and number of columns
[x y z]=size(I);
L=y;
R=zeros(x,y,z);
R(:,L+1:y,:)=I(:,1:y-L,:);
R(:,1:L,:)=I(:,L:-1:1,:);
R=uint8(R);
end
```

A.2.2- Find max of vertical (Section 4.1.5):

```
function [maxvertical]=maxverticals(image)
[Yy,Xx]=find(image);
lines=zeros(15,2);
N=length(Xx);
for i=1:N;
   for j=1:N;
      if (abs(Xx(i)-Xx(j)))<=4
         Xx(j)=round((Xx(i)+Xx(j))/2);
         Xx(i)=Xx(j);
      end
   end
end
E = countelement(Xx);
```

```
j=1;
for i=1:size(E,1)
   if E(i,2)>=35;
      lines(j,1)=E(i,1);lines(j,2)=E(i,2); j=j+1;
   end
end
L=length(find(lines(:,1)));
for i=1:L
   k=1;
   for j=1:N
      if Xx(j)==lines(i,1)
      points(i,k)=Yy(j); k=k+1;
      end
   end
end
B=sort(points,2); %get highest points in the first column
for i=1:L
   I=find(B(i,:), 1, 'first');
   maxvertical(i,2)=B(i,I);
end
maxvertical(:,1)=lines(1:L,1);
maxvertical=sortrows(maxvertical,2);
if size(maxvertical,1)>8
   maxvertical=maxvertical(1:8,:);
end
maxvertical=sortrows(maxvertical,1)
```

A.2.3- Find middle of horizontal (Section 4.16):

```
function [maxhoriz]=horizontals(image,p)
image=image(p(2)-65:p(2)+50,p(1)-250:p(1)+280);
[Yy,Xx]=find(image);
A=[Yy,Xx];
A=sortrows(A,1);
```

```
Yy=A(:,1);
Xx=A(:,2);
m=size(image,2);
N=length(Yy);
for i=1:N-1;
   for j=i+1:N;
      if ((abs(Yy(i)-Yy(j)))<=2 && (abs(Xx(i)-Xx(j)))<=15)
         Yy(j)=min(Yy(i),Yy(j));
      end
   end
end
A=[Yy,Xx];
A=sortrows(A,1);
 image(:,:)=0;
 for i = 1:size(A,1)
 image(A(i,1),A(i,2)) = 1;
 end
Yy=A(:,1);
Xx=A(:,2);
E = countelement(Yy);
l=1;
EE=zeros(10,2);
for i=1:size(E,1)
   if E(i,2)>5 && E(i,2)<50
      EE(l,:)=[E(i,1) E(i,2)];l=l+1;
   end
end
EEy=cell(1,size(EE,1));
for i=1:size(EE,1)
   EEy{1,i}=zeros(size(EE,2),2);
   k=1;
   for j=1:N
      if Yy(j)==EE(i,1)
         EEy{1,i}(k,:)=[Yy(j) Xx(j)];k=k+1;
```

```
        end
    end
end
for i=1:size(EEy,2)
    EEy{i}=sortrows(EEy{i},2);
end
points=zeros(7,2);
l=1;
for i=1:size(EEy,2)
        k=1;
        while k+5 <= size(EEy{i},1)
            a=EEy{i}(k+5,2)-EEy{i}(k,2);
            if a<=12
                points(l,:)=[round(mean([EEy{i}(k,2),EEy{i}(k+5,2)]))+2, EEy{i}(k,1)];
                l=l+1;
            end
             k=k+find((EEy{i}(k+1:size(EEy{i},1),2))> (EEy{i}(k,2)+10),1,'first');
        end
end
for i=1:size(points,1)
    if points(i,2)~=0
    points(i,2)=points(i,2)+p(2)-65;
    end
end
for i=1:size(points,1)
    if points(i,1)~=0
    points(i,1)=points(i,1)+p(1)-250;
    end
end
maxhoriz=points;
end
```

A.2.4- Find End of teats in the thermal image (Section 4.1.7):

```
function [teats,hiddenteats]=endteats(maxhoriz)
```

```
teats=zeros(2,2);
max=find(maxhoriz(:,1),1,'last');
k=1;
maxhoriz2=zeros(max,2);
for i=1:size(maxhoriz,1)
   if maxhoriz(i,1)~=0
      maxhoriz2(k,:)=maxhoriz(i,:);
      k=k+1;
   end
end
 m=round(mean(maxhoriz2(:,2)));
maxhoriz2=sortrows(maxhoriz2,1);
i=1;
k=1;
while k==1
   if abs(maxhoriz2(i,2)-m)<=30
      teats(k,:)=maxhoriz2(i,:);
      k=k+1;
   end
      i=i+1;
end
 while k<=4 && i<=max
   if abs(maxhoriz2(i,1)-teats(k-1,1))>30 && abs(maxhoriz2(i,2)-m)<=30
      teats(k,:)=maxhoriz2(i,:);
      k=k+1;
   end
   i=i+1;
end
hiddenteats=abs(k-5);
end
```

A.3- Apply homography:

```
function [teats_left,teats_right]=getteatsOptical(teats,Hright,Hleft,left,right)
teats=teats';
teats(3,:)=1;
teats_r=Hright*teats;
teats_l=Hleft*teats;
teats_right=teats_r(1:2,:);
teats_left=teats_l(1:2,:);
 end
```

A.4- Find end of teat in optical image:

```
function [imageopt,BW,teat]=locatinoptical(opimage,teats_p,i)
teats_p=round(teats_p);
imageopt=opimage(teats_p(2,i)-20:teats_p(2,i)+20,teats_p(1,i)-20:teats_p(1,i)+20);
SE=ones(9,10);
SE(1:2,1)=0; SE(1:2,10)=0; SE(1,2)=0; SE(1,9)=0;
BWtest=imdilate(imageopt,SE);
BW1=im2bw(BWtest,0.3);
SEr = strel('rectangle', [10 7]);
BW=imdilate(BW1,SEr);
if size(find(BW),1)>= 1100
   teat=locateteat(BW);
      teat(1,1)=teat(1,1);  %teat(1,1)=teat(1,1)+teats_p(1,i)-20;
      teat(1,2)=teat(1,2);  %teat(1,2)=teat(1,2)+teats_p(2,i)-20;
else
      teat(1,1)=20;  %teat(1,1)=teats_p(1,i);
      teat(1,2)=20;  %teat(1,2)=teats_p(2,i);
end
end

 %%%%%%%%%%%%%%%%%%%%%%%%
function [teat]=locateteat(BW)
ss=size(find(~BW),1); % get number of black pixels
```

```
BW=~BW;            %
[Yy,Xx]=find(BW);   %   coordonnes of objects
F=[Yy,Xx];         %
F=sortrows(F,1);    %    sort them horizontally
Fy=F(:,1);         %
E= countelement(Fy);%     count number of black pixels by row
s=size(BW,2);       %
i=1;               %
getthe2points=false;%       initialisation of booleen variable
while(getthe2points==0)&&(i<size(E,1))                    %
   if E(i,2)>9                                        % initialise first point
      ind=find(BW(E(i,1),:),1,'first');                    %
      while(ind+10<s+1)&& (getthe2points==0)                %
         if BW(E(i,1),ind:ind+9)==1                      % verify the row
           getthe2points=true;                          %
         else
            ind=ind+find(BW(E(i,1),ind+1:40)==0,1,'first')+1;   %  get two points
         end
      end
      if getthe2points==0                                %
         i=i+1;                                       %
      end
   else
      i=i+1;                                          %
   end
end
if getthe2points                                       %
   if find(BW(E(i,1),ind:s)==0,1,'first')                  %
     ind2=ind-1+find(BW(i,ind:s)==0,1,'first');            %
   else
      ind2=s;                                         %
   end
end
```

```
   if ss<1400 && abs(ind2-ind)>=24
      teat(1,1)=round((abs(ind2-ind)/4)+ind);                    %
      teat(1,2)=round(E(i,1)-15);
   else
      teat(1,1)=round((abs(ind2-ind)/2)+ind);                    %
      teat(1,2)=round(E(i,1)-15);
   end
end
```

A.5- Stereo triangulation:

The code of this program is the same as the one provided by (Bouguet, 2008b) with small changes introduced to the output of the function. The output result is given in the frame system related to the right camera.

```
function [XR] = stereo_triangulation(xL,xR,om,T,fc_left,cc_left,kc_left,alpha_c_left,fc_right,cc_right,kc_right,alpha_c_right)
% [XL,XR] = stereo_triangulation(xL,xR,om,T,fc_left,cc_left,kc_left,alpha_c_left,fc_right,cc_right,kc_right,alpha_c_right),
% Function that computes the position of a set on N points given the left and right image projections.
% The cameras are assumed to be calibrated, intrinsically, and extrinsically.
% Input:
%       xL: 2xN matrix of pixel coordinates in the left image
%       xR: 2xN matrix of pixel coordinates in the right image
%          om,T: rotation vector and translation vector between right and left cameras (output of stereo calibration)
%           fc_left,cc_left,...: intrinsic parameters of the left camera  (output of stereo calibration)
%          fc_right,cc_right,...: intrinsic parameters of the right camera (output of stereo calibration)
% Output:
%       XL: 3xN matrix of coordinates of the points in the left camera reference frame
```

```
%       XR: 3xN matrix of coordinates of the points in the right camera reference frame
% Note: XR and XL are related to each other through the rigid motion equation: XR = R * XL + T, where R = rodrigues(om)
%       For more information, visit http://www.vision.caltech.edu/bouguetj/calib_doc/htmls/example5.html
% (c) Jean-Yves Bouguet - Intel Corporation - April 9th, 2003
%--- Normalize the image projection according to the intrinsic parameters of the left and right cameras
xt = normalize_pixel(xL,fc_left,cc_left,kc_left,alpha_c_left);
xtt = normalize_pixel(xR,fc_right,cc_right,kc_right,alpha_c_right);
%--- Extend the normalized projections in homogeneous coordinates
xt = [xt;ones(1,size(xt,2))];
xtt = [xtt;ones(1,size(xtt,2))];
%--- Number of points:
N = size(xt,2);
%--- Rotation matrix corresponding to the rigid motion between left and right cameras:
R = rodrigues(om);
%--- Triangulation of the rays in 3D space:
u = R * xt;
n_xt2 = dot(xt,xt);
n_xtt2 = dot(xtt,xtt);
T_vect = repmat(T, [1 N]);
DD = n_xt2 .* n_xtt2 - dot(u,xtt).^2;
dot_uT = dot(u,T_vect);
dot_xttT = dot(xtt,T_vect);
dot_xttu = dot(u,xtt);
NN1 = dot_xttu.*dot_xttT - n_xtt2 .* dot_uT;
NN2 = n_xt2.*dot_xttT - dot_uT.*dot_xttu;
Zt = NN1./DD;
Ztt = NN2./DD;
X1 = xt .* repmat(Zt,[3 1]);
X2 = R'*(xtt.*repmat(Ztt,[3,1])  - T_vect);
%--- Right coordinates:
XR = (R*XL + T_vect)';
```