Close-range videometry -

Design and calibration of a mono-camera system for dynamic purposes.

Ivar Maalen-Johansen
Dept. of surveying
Agricultural Univ. of Norway

WG V/1

Abstract.

The demand for measurements of a moving surface inside a small vacuum-chamber is meet by designing a special videometric system described in the paper.
Because of initial space restrictions, one miniature CCD-camera together with mirrors form a split-image system. Calibration of this system includes additional parameters to describe the distortion of the optical system.
Reduced radiometric resolution and use of only one video-field makes it possible to register a sufficient number of subsequent images in the internal buffer of the framegrabber. The effects of the special solutions on the accuracy are regarded in order to meet the accuracy requirements.

Keywords:

Videometry, close-range, split-image, calibration, system design.

1. Introduction.

The task of studying the condition of cow-teats during an automated milking process is desirable from an agricultural researchers point of view to optimize the process. Today the only possible insight into the process is a manual examination of the teats before and after milking.
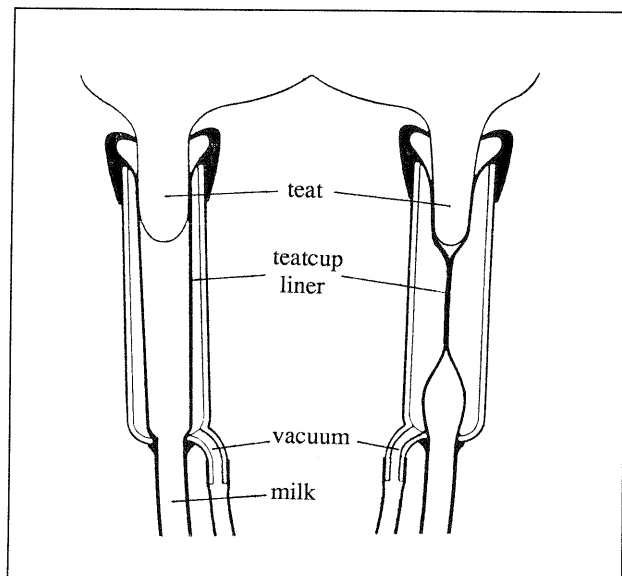

Figure 1. Overview of the original milking machine.

The aim of our project was to determine some unknown parameters of the milking process, such as distribution of the pressure from the teatcup liner on the teat.

It is vital to the success of the milking process that the impact from the teatcup liner on the teat is correct. The pressure from the liner is to maintain the initial shape during the massage phase. Among the consequences of defective pressure are slower milking and risk of udder disease.

To estimate the forces between liner and teat it is necessary to know the shape of the liner through the cycle. A number of surface-models have to be measured without disturbing the function of the teatcup liner inside the vacuum-chamber. Therefore no sensors can be fixed directly to the liner.

In addition, it should be possible to collect the measurements in the cow barn environment.

These reasons led to a proposal to make use of the remote sensing features of digital photogrammetry. A cooperation between two departments at the Agricultural University, Dept. of Surveying and Dept. of Agricultural Engineering, resulted in the videometry system described in this article.

We have emphasized the calibration of our system since we are using both an "off-the-shelf" CCD-camera with no high-grade optics and have included mirrors in the imaging system. This, besides the sources of error in the CCD-technology, makes the camera modelling and calibration fundamental tasks in a system build-up.

The calibration used in this system is a on-the-job calibration where a calibration frame is imaged in the place of the teatcup liner.

2. Hardware of the system.

When designing the system, the hardware was given priority because of the special requirements of the system.

To make it possible to measure the teatcup liner inside the milking machine without disturbing the process, an adapter with a window to the teatcup had to be designed. The vacuum process should remain intact while measurements are taken. Limited space and limited budget were other constraints in the system consideration.

The hardware solution was a mono-camera system with a mirror arrangement in front of the camera to create stereoscopic split images. In general, one-camera designs have some obvious advantages in price, calibration, etc. (El-Hakim 1990), and in this set-up factors like size, data quantity and synchronization also play decisive roles.

The aim was also to use a "bottom to top" strategy in the design and start with uncomplicated solutions.

The selected camera has a remote miniature head sized 17x55 mm with resolution better than 360 (horizontal) x 420 (vertical). Since the simultaneous images are recorded on the same frame, no synchronization or pairing problems occur.

Figure 2 shows two cross-sections of the teatcup and the camera adapter.

The glass plate separate the mirrors and camera from the teatcup where varying vacuum occurs. Illumination is by external light sources through a diffusive window to avoid disturbing reflections. The prototype adapter is made of brass, but, the next version will be of a lighter composite material to keep down the additional weight of the adapter.

3. Functions of the system.

To get the required results of the system, solutions from different parts of the subject area had to be adapted and adjusted to the tasks. The accuracy requirements of the results were the most important factors for the choices of methods of measurement, data handling, estimation, etc. The system overview is shown by data flow diagram (DFD) (Gane and Sarson, 1979) in figure 3, and the more detailed decompositions are shown in the following subsections.

In this article the processes of calibration and measurement are emphasized, both because they are the two most decisive parts of any system and because of the special solutions in our system.
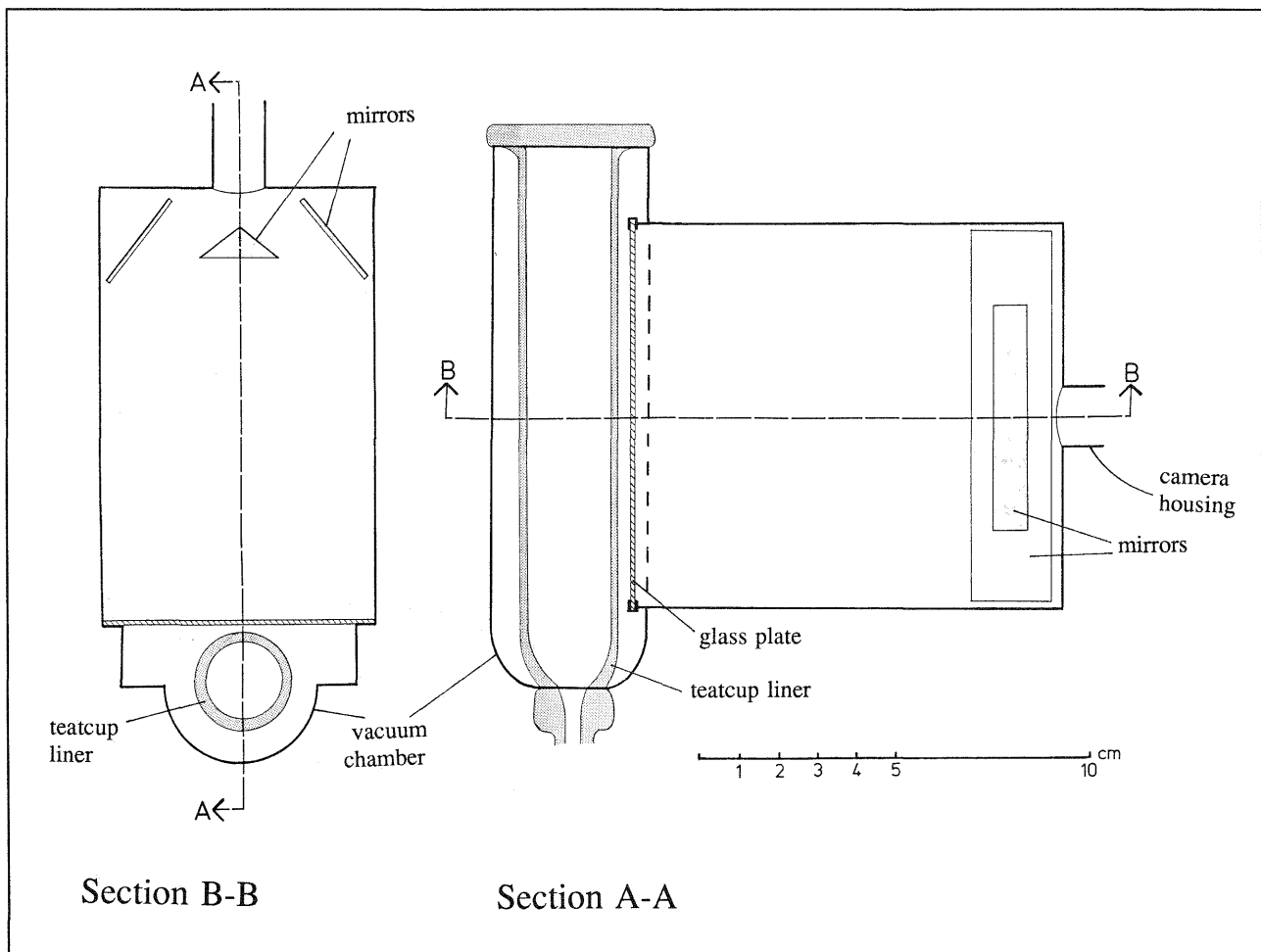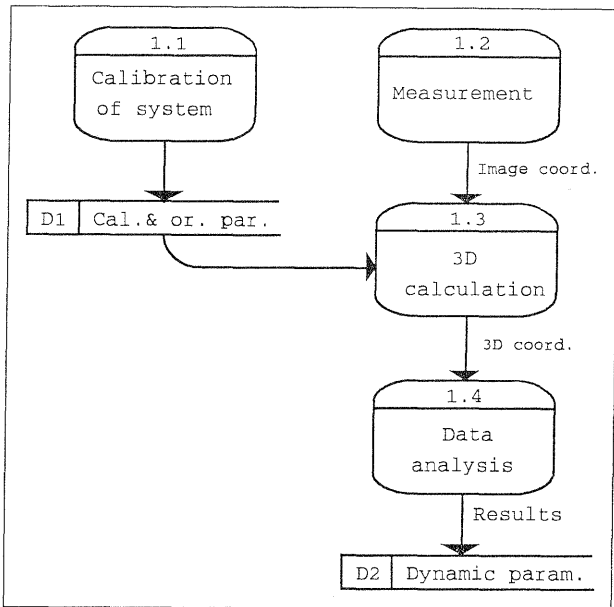

Figure 2. Teatcup with adapter.

Figure 3. First order decomposition of the videometry system.

3.1 Measurement.

The aim of the measurement process is to obtain image coordinates of the targets on the object for further use in the object reconstruction. The choice of measurement method and parameters determine the accuracy of the system and must be regarded specially.
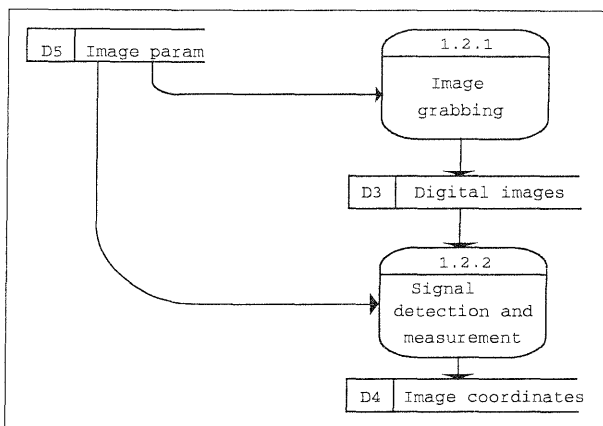

Figure 4. Decomposition of the measurement process.

3.2 Targeting.

In most close-range applications artificial targeting is required to produce sufficient quality and number of data points. The need for artificial targeting was obvious for the teatcup liner application since the surface of the liners are smooth and uniform and do not provide any photographic contrast.

Two types of target points were considered for the teatcup liners; both retroflective targets and projection of targets onto the surface. When choosing the targeting it has to be considered how the targets are imaged in the camera and the possibilities of having homologous data points in the following calculations. Other important criteria for the decision were speed of execution and simplicity of procedure and hardware.

The use of retroflective targets is common in close-range use. Circular targets of appropriate size are fixed to the surface in a suitable pattern. This method ensures symmetric targets and a best possible conformal representation of all the targets in the image. This also forms the basis of successful target centre determination.

Different methods of targeting with structured light were considered. Structured light can be used both in an active oriented sensor (active triangulation) and for target generation (passive triangulation). A concept of regarding a raster projector as an additional camera in a rasterstereographic system is described by Frobin and Hierholzer (1981) among others. The method has some advantages over conventional close range methods, according to Ethrog (1991): a) Synchronization of two cameras is not needed; b) Calibration of the projector with the raster diapositive need only be done once; and c) When using a linear pattern in the projection, the method is well suited for automated evaluation.
Another variant of triangulation is Moiré technique based on multistripe pattern projection (Tiziani, 1989).

Projection of structured light (i.e., laser) in either a stochastic pattern or a regular raster, is another method of targeting. In this method the surface is imaged simultaneously in two or more images by usual photogrammetric procedure. For obvious reasons, scanning arrangements are unsuitable for moving surfaces.

The target projecting methods have the advantage of being totally contact-free and not interfering with the measured object. These types of targeting do not ensure homologue points in subsequent images since there is no direct reference to the object.

One precondition for success with such targeting is a bright model surface. The rubber surface of the teatcup liner is dull dark and made such targeting unfit for our case.

The choice for our prototype system was a regular pattern of circular retroflective targets sized ø 1.5 mm. with intervals of 6.5 mm. This target size corresponds to approximately 5 pixels in the image, which seems to be a favorable size for passive targets (Trinder 1989).
Although this method implies targets fixed onto the surface, we expect no influence from the targets on the characteristics of the teatcup liner. We also obtain directly comparable results, e.g. tension in the liner, from the measured target intervals.
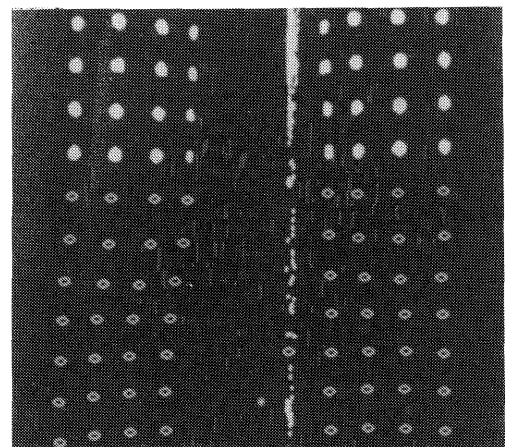

Figure 5. A slightly flattened teatcup liner imaged in the left and right image. The ellipses indicate located targets.

3.3 Image acquisition.

The image acquisition process is the sequence from the imaging of the object on the sensor to the stored digital image. This can be described as in figure 6.
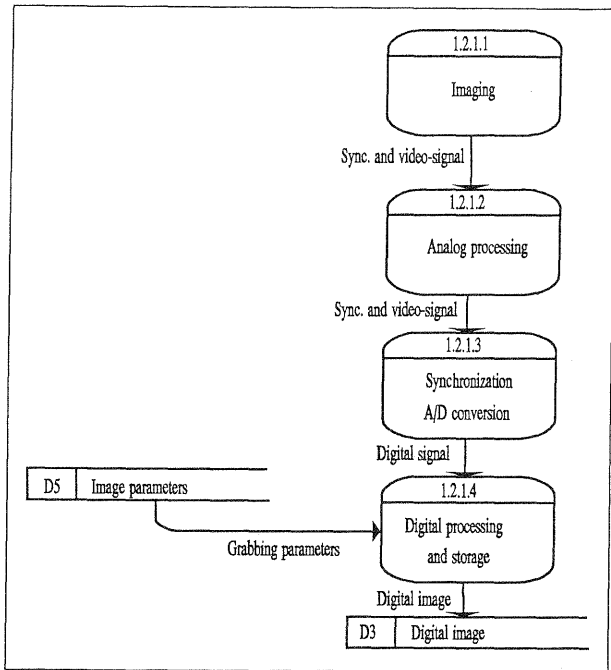

Figure 6. Decomposition of the image acquisition process.

The ordinary frame rate of CCD-cameras (CCIR, European standard) is 25 Hz, i.e., 25 images are captured per second. Without interlacing it is possible to use only one of the two videofields to gain a doubled rate. In measurements of fast moving objects this is a quite common method (e.g., Baltsavias and Stallmann, 1990).

Different methods are used to interpolate the 'missing' field in order to maintain complete frames for the later analysis.
The aim of this process is to maintain the spatial resolution in the y-direction without significant loss of precision.

The method applied in our system is based on a convolution performing a filtering of the field by a 3x3 kernel. This process also eliminates discontinuities in the line-direction (x-direction). Another closely related approach is linear interpolation in the y-direction. In a future version we will keep the field intact (without interpolation) throughout the measurement process in order to control and increase the precision and accuracy.

3.4 Reduced radiometric resolution of the image.

The analog video-signal from the CCD-sensor are usually converted into an 8 bit-number ranging from 0 (black) to 255 (white) using an analog to digital (A/D) converter. The output is stored in a frame buffer on the frame grabber board and afterwards read out to the host computer. Using a reduced radiometric resolution, i.e., a lower number than 8, the number of frames that can be stored in the buffer is increased. When the read-out from the frame buffer to the computer is slower than the video-rate, it is favorable to have a required storage capacity in the frame buffer, especially when dealing with fast moving dynamic scenes.

The influence of the number of quantization levels on the precision and accuracy of pointing is investigated by Trinder (1989) showing no significant deterioration for quantization above 4 bits/pixel. Below 5 bit/pixel the pointing precision to circular targets decrease considerably. In a binary image (1 bit/pixel) the pointing precision is estimated to be approximately 10-15 times larger than with 5 bits/pixel, or 15-20 times larger than with full quantization (8 bits/pixel).

Similar results are obtained by applying the 'locale' concept (Havelock, 1989 and 1991) where the sizes of regions of indistinguishable object position are the basis for the estimation of precision. As an example it is stated that an estimate of the position of a small circular target in a binary image has a precision of up to 0.3 pixel in the worst case. With an image scale of 1:30 this indicates a precision better than 0.1 mm. on the object, which is well inside our requirement. This allows storage of a continuous sequence of 32 binary images in the frame buffer at the video rate.

3.5 Target location.

The process of detection and measurement of the position of the targets in digital images requires subpixel resolution because the data is a digital representation of an analog signal sampled onto a discrete array and simultaneously quantised to a finite number of levels. To obtain a satisfactory measurement accuracy it is necessary to measure in between the sample positions.

The precision of the determination depends upon the method, image quality, quantization levels, pixel size and noise.
Since circular shaped targets were chosen in our system, methods suitable for location of such targets are of main interest here.

Different techniques for subpixel location fall into the categories of interpolation, correlation, centroiding, edge analysis or shape-based methods. The performance of different methods is investigated with respect to spatial and radiometric resolution and accuracy in West and Clarke (1990). Use was made of both simulated data and real data from optical triangulation with 1D sensors and laser light sources. Three of the categories were found to be the most applicable to this type of task: interpolation, correlation and centroiding. The results show that most techniques can perform accuracy better than 0.1 pixel. A weighted centroid method obtained best results in simulation while a Vernier method (Tian and Huhns, 1986) was better on real data.

For use with circular targets, or symmetric targets in general, variants of the centroid method are the most common. The techniques differ in the way the centroid is computed and the pixel values used. In the most simple approach the standard first order moment is computed using the grey values of the target in the image. Thresholding is used to reduce the number of pixels in the computation. With a symmetric object the centroid will give a perfect result. Asymmetry together with noise and quantisation are the main contributors to the loss of accuracy (West and Clarke, 1990).
The centroid method is based on formula (1) and (2):

$$x = \frac{\sum i \cdot p_{ij}}{\sum p_{ij}} \tag{1}$$

$$y = \frac{\sum j \cdot p_{ij}}{\sum p_{ij}} \qquad (2)$$

where p_{ij} is the intensity value of the pixel located in row i and column j.

Asymmetric effects can be reduced by adding a weighting factor equal to the intensity value of the pixel. Also a correct threshold value for the particular target size and level of asymmetry will eliminate the systematic errors (Trinder, 1989). The threshold value T is estimated from:

$$T = 74 \ (SF)^{1.3} \ A^{-1} \qquad (3)$$

where SF is the 2σ-width of the spread function and A the target size.

In the same investigation it is stated that if the targets are larger than 4x the 2σ width of the spread function and the correct threshold value is used, the systematic errors of target pointing will be less than 0.02 pixel. This result corresponds to an 8 bits/pixel quantisation and the reduced resolution in our system will deteriorate the results as mentioned earlier.

To investigate the precision of the target location in our system a sequence of 32 measurements of the 108 targets on the calibration frame were taken. The results indicate a pointing precision of 0.170 pixel in x-direction and 0.225 pixel in y-direction. A possible explanation of the larger deviations in y is the reduced resolution in the y-direction because only one video-field is used. Otherwise the deviation in x should be the largest because of line-jittering.

3.6 Calibration.

The task of the calibration of the system is to estimate the interior and exterior orientation of the camera. Since we are doing an on-the-job calibration, both orientations can be performed in the same adjustment. We must consider a system as consisting of two cameras with imaginary positions and rotations because of the mirrors and split-image arrangements. Figure 7 shows the geometry of the imaging system.

The two halves of the sensor form the two images of the imaginary cameras. The interior orientation denotes the estimation of the parameters in the functional camera model. The functional model of the system describes the imaging of an object point past the mirrors, through the lens with its distortions and onto the sensor. The accuracy of the system depends strongly upon the model chosen and how well this model suites the real conditions. Kilpelä (1981) proposes a number of different parameter sets to be used to extend the pinhole model given by the perspective equations. Our formulation of the functional model is a variant of the "physical model" from that paper extended with parameters to help describe the mirror distortions. Our model then includes additional parameters for affinity, lack of orthogonality, radial and decentering distortions in addition to mirror distortions.

Another possible method used to describe the influence of the mirrors is the use of mathematical surfaces, i.e. B-splines (Åmdal, et.al., 1990). This would possibly model the local anomalis in a better way, but require a larger set of datapoints in our calibration.
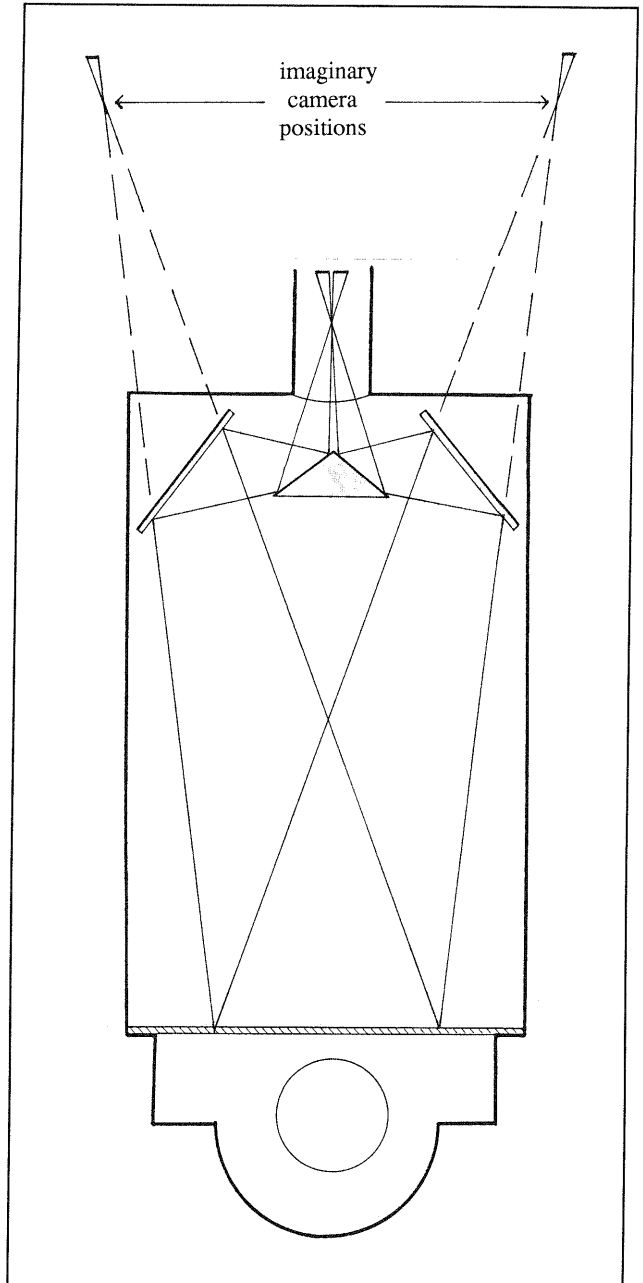

Figure 7. Geometry of the imaging system.

The functional model used in our system is given in equations 4 to 7, where:

x,y	: Observed image coordinates.
x_o, y_o	: Center of the frame memory.
x_c, y_c	: Undistorted image coordinates, corrected for distortion effects.
X,Y,Z	: Object coordinates of an imaged point.
X_o, Y_o, Z_o	: Object coordinates of the origin (projection center) of the camera.
f	: Focal length of the camera.
r	: Distance from (x_o, y_o) to (x_c, y_c).
r_{ij}	: Elements of the rotation matrix.
$b_1, ..., b_9$	: Parameters describing the distortions.

490

$$x = x_0 - f \frac{r_{11}(X-X_0)+r_{12}(Y-Y_0)+r_{13}(Z-Z_0)}{r_{31}(X-X_0)+r_{32}(Y-Y_0)+r_{33}(Z-Z_0)} + dx \qquad (4)$$

$$y = y_0 - f \frac{r_{21}(X-X_0)+r_{22}(Y-Y_0)+r_{23}(Z-Z_0)}{r_{31}(X-X_0)+r_{32}(Y-Y_0)+r_{33}(Z-Z_0)} + dy \qquad (5)$$

$$dx = -b_1 x + b_2 y + b_3 x_c r^2 + b_4 x_c r^4 + b_5 x_c r^6$$
$$+ b_6(r^2 + 2x_c^2) + 2b_7 x_c y_c - b_8 x_c y_c^2 \qquad (6)$$

$$dy = -b_1 y + b_2 x + b_3 y_c r^2 + b_4 y_c r^4 + b_5 y_c r^6$$
$$+ 2b_6 x_c y_c + b_7(r^2 + 2y_c^2) - b_9 x_c^2 y_c \qquad (7)$$

The first steps involved in the calibration process are similar to the procedure for the measurements. The process of calibration is described in figure 8.
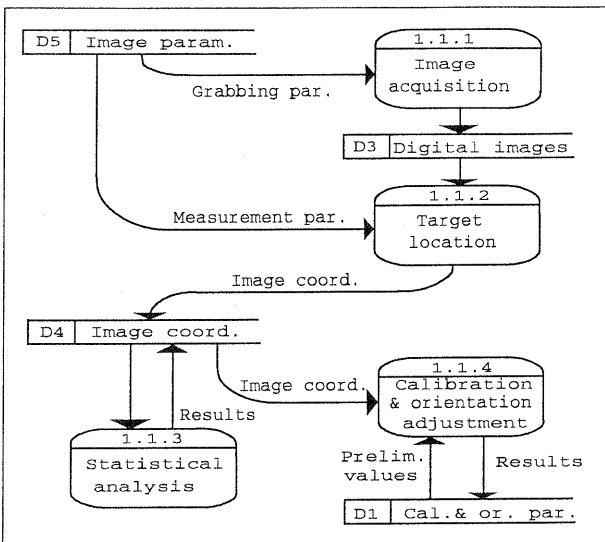
Figure 8. Decomposition of the calibration process.

In order to eliminate effects of syncronization of the videosignal (i.e. line-jitter) which are not investigated here, a sequence of images of the calibration frame is captured. The target location is established as described earlier and the mean values of the image coordinates are used in the following adjustment as observations. The stochastic model used in this investigation assume the observations to be uncorrelated and having a normal distribution with equal variance. These are truly only assumptions, but common practice in similar investigations.

The on-the-job calibration is performed by placing a calibration frame with accurately determined targets in the measuring space. This frame is targeted as described in 3.2 and the geometry is shown in figure 9.

The programme used is a least squares bundle adjustment adapted to our task. Both interior and exterior orientation can be estimated and the desired number of additional parameters can be included in the mathematical model. Only parameters being significantly different from zero remain in the model throughout the adjustment.

Results from our system show considerable differences between the two imaginary cameras. This can be explained by different optical conditions for the two cameras, concerning both the mirror arrangement and the two halves of the camera optics. Further investigations of other additional parameters must be made if higher accuracy is required from the system.
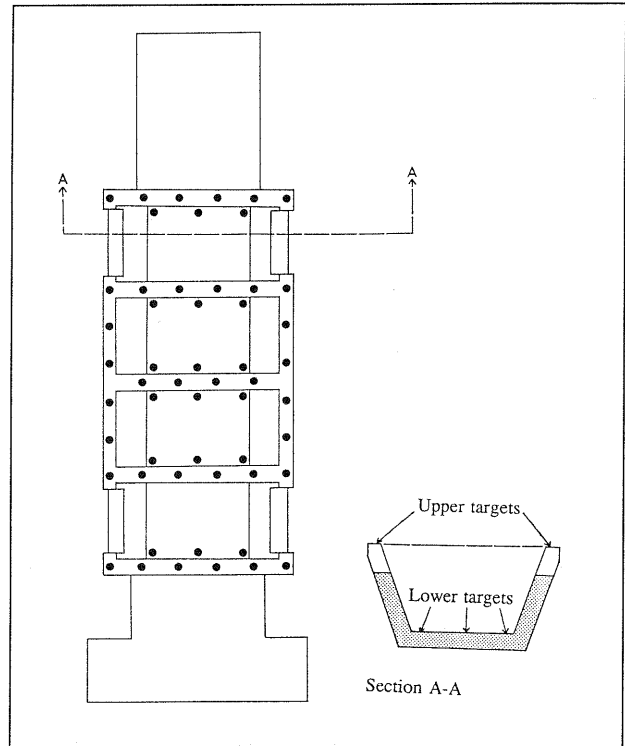
Figure 9. Calibration frame.

3.7 3D-model calculation.

The observed image coordinates must be corrected because the functional model of the imaging system is not a pinhole model. Mirror and lens distortion, affinity, lack of orthogonality and the center of the frame memory are parameters estimated in the calibration that form the transformation between observed and the corrected image coordinates.

One central problem is finding corresponding points (targets) in the two (or more) images. Using targets of identical size and shape, methods like feature based image matching (Förstner,1986) are not the most suitable. Introducing topological constraints like epipolar geometry will be a more effective approach here. Because of prior knowledge of the object geometry the identification of the targets in the images can be performed by using the perspective equations between object and images.
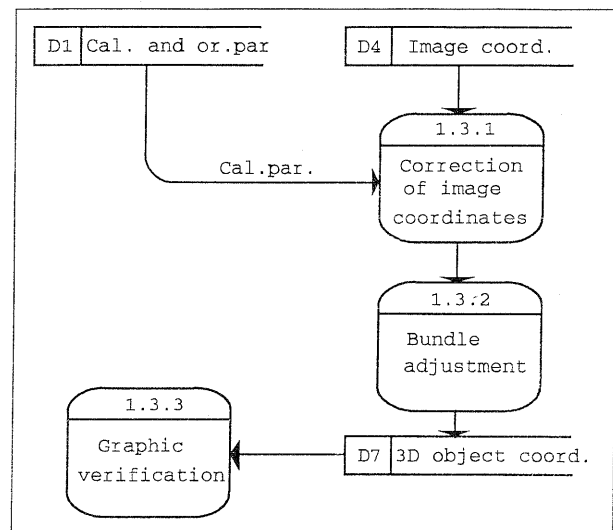
Figure 10. Decomposition of 3D model calculation.

The 3D reconstruction is a bundle adjustment method performing the transformation between the image coordinate system and the object coordinate system. This is an intersection using the exterior parameters also obtained in the previous calibration.

To obtain a representative measure of accuracy of the system a targeted object is imaged inside the actual area. Calculated coordinates are compared to the given values and the deviations are used to estimate the accuracy. The given values are found by accurate photogrammetric methods.

The precision of measured coordinates is estimated from repeated measurements.

Results of this analysis are given in table 1.

	X	Y	Z	XYZ
Accuracy (RMSE)	0.056	0.101	0.082	0.141
Precision (σ)	0.051	0.035	0.135	0.148

Table 1. Results from accuracy investigation (all in mm.).

Different methods can be applied to verify the 3D model. In our set-up, 3D graphics software was used to generate grid models for visual verification. A more sophisticated approach is surface modelling by generating synthetic images (Thune, 1991). Perspective views can be generated from the estimated object model using the pixel values and the camera model in a simple ray casting technique.

3.8 Data analysis.

The initial task of our system was to obtain characteristic parameters of the functioning of the teatcup liner. Using 3D object models from different phases of the milking process, it is possible to estimate variable parameters like shape and size of the surface. From these the required data of tension in the liner and the distribution of the pressure from the liner on the teat can be obtained by means of a physical model. This work is performed by the Dept. of Agricultural Engineering at AUN. An example of result from this analysis is given in figure 11.

4. Conclusions

This paper descibes the work of designing a videometric system to meet the special requirements of measuring a moving surface inside a small vacuum chamber.

The work consists of a system component description and an analysis of each component regarding mainly their influence on the overall accuracy.

Binary radiometric resolution of the images and use of only one videoframe are the main contributors to the accuracy loss. These choices were made in order to handle the dataflow from the fast moving dynamic scenes and still be inside the accuracy requirements.

Results from the calibration of the system with both mirrors and CCD-camera indicates considerable local distortions in the images. These were probably caused by low-cost optics and mirrors which were impossible to model by global parameters.

Both the resulting spatial accuracy and precision of the system were inside 2/10 mm and satisfied the requirements.
When higher accuracy standards are required other choices can be made and another more sophisticated mathematical model used.
The data provides for a new analysis of the dynamics of the teatcup liner, and, the results are promising so far.
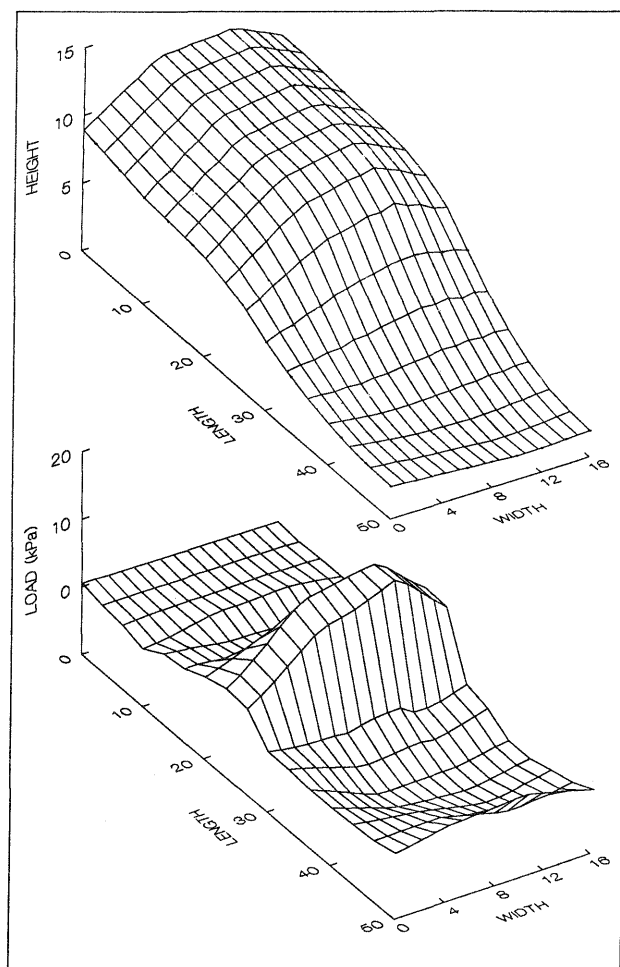

Figure 11. Estimated 3D model of the teatcup liner (upper plot) and pressure from the liner on the teat (lower plot).

5. Acknowledgements

6. References

Baltsavias, E.P., and Stallmann, D., 1990. Trinocular vision for automatic and robust 3-D determination of the trajectories of moving objects. In: Close-Range Photogrammetry Meets Machine Vision, A.Gruen, E.P. Baltsavias, Editors, Proc.SPIE 1395, pp.620-629.

El-Hakim, S.F., 1990. Some solutions to vision dimensional metrology problems. In: Close-Range Photogrammetry Meets Machine Vision, A.Gruen, E.P. Baltsavias, Editors, Proc.SPIE 1395, pp.480-487.

Ethrog, U., 1991. Rasterstereography and double-rasterstereography: Two methods for measuring fast dynamic changes of laboratory model surfaces. J. Photogramm. Remote Sensing 46, pp.269-282.

Frobin, W., and Hierholzer, E., 1981. Rasterstereography: A photogrammetric method for measurement of body surfaces. Photogrammetric Engineering and Remote Sensing 47(12) pp.1717-1724.

Förstner, W., 1986. A Feature Based Correspondence Algorithm for Image Mathcing. Int. Arch. of Photogrammetry and Remote Sensing 26-III, pp.1-17.

Gane, C., and Sarson, T., 1979. Structural Systems Analyses: tools and techniques. Prentice-Hall Inc, pp.8-24.

Havelock, D.I.,1989. Geometric precision in noise-free digital images. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.PAMI-11, No.10, pp.1065-1075.

Havelock, D.I.,1991. The Topology of Locales and Its Effects on Posision Uncertainty. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.PAMI-13, No.4, pp.380-386.

Kilpelä, E., 1981.Compensation of systematic errors of image and model coordinates. Photogrammetria 37, pp.1-13.

Thune, N., 1991. Stereo Vision - An Integrated Approach. Ph.D. dissertation, Division of Computer Systems and Telematics, The Norwegian Institute of Technology, pp.110-113.

Tian, Q., and Huhns, M.N., 1986.Algorithms for subpixel registration. Computer Vision, Graphics and Image processing, Vol.35, pp.220-233.

Tiziani, H., 1989. Optical 3D Measurement Techniques - a Survey. In: Optical 3D Measurement Techniques, A. Gruen, and H. Kahmen, Editors, H. Wichmann Verlag, pp.3-21.

Trinder, J.C., 1989. Precision of Digital Target Location. Photogrammetric Engineering and Remote Sensing 55(6) pp.883-886.

West, G.A.W., and Clarke, T.A., 1990. A survey and examination of subpixel measurement techniques. In: Close-Range Photogrammetry Meets Machine Vision, A.Gruen, E.P. Baltsavias, Editors, Proc.SPIE 1395, pp.456-463.

Åmdal, K., Thune, N., and Dørum, O.H., 1990. Correction of Image Deformations using B-Spline Surfaces in Digital Close Range Photogrammetry. In: Close-Range Photogrammetry Meets Machine Vision, A.Gruen, E.P. Baltsavias, Editors, Proc.SPIE 1395, pp.845-854.