Machine Vision and Applications (1995) 8:187–193

Machine Vision and Applications

Segmentation and tracking of piglets in images

N.J.B. McFarlane and C.P. Schofield

Silsoe Research Institute, Wrest Park, Silsoe, Bedford, MK45 4HS, UK

Abstract. An algorithm was developed for the segmentation and tracking of piglets and tested on a 200-image sequence of 10 piglets moving on a straw background. The image-capture rate was 1 image/140 ms. The segmentation method was a combination of image differencing with respect to a median background and a Laplacian operator. The features tracked were blob edges in the segmented image. During tracking, the piglets were modelled as ellipses initialised on the blobs. Each piglet was tracked by searching for blob edges in an elliptical window about the piglet's position, which was predicted from its previous two positions.

Key words: Tracking – Segmentation – Pigs – Animals – Computer vision

1 Introduction

The tracking of animals by machine vision has many possible applications in surveillance and livestock monitoring. One such application is the monitoring of newly born piglets in farrowing pens. Sows frequently give birth to their litters with little or no attendance from the stockman, who may have several hundred pigs in his charge. An automatic monitoring system, capable of alerting the stockman to trouble during farrowing, such as an unduly long period between piglets being born, or piglets being unable to find the sow's teats, would therefore be beneficial in terms of both productivity and animal welfare.

There are many other applications in which the automatic location and tracking of animals by image analysis will be beneficial to their health, welfare and productivity (Deschazer et al. 1988; Schofield 1993). Monitoring the location of pigs within a pen over a period of time provides information relevant to their activity and their feeding, drinking and resting behaviour. (Van der Stuyft et al. 1991). The ability to track pig movements is an important step towards developing systems to monitor their behaviour and condition. This provides the stockman with behavioural information such as changes in health, activity, fighting, bullying, and tail biting, which he can act upon as required. Activity and resting behaviour are indicative of heat stress caused by poor ventilation patterns and inaccurate temperature control (Geers et al. 1990). Huddling for warmth indicates high heat loss and has the effect of increasing physical activity, resulting in lower feed conversion (Mount 1968). Recognition of these indicators through automatic image analysis allows the behaviour of pigs to be used to control their environment.

The tracking of piglets is difficult because, as with most animate objects, their movement is subject to unpredictable manoeuvres generated by the animals themselves, the features which distinguish them from other objects in the background may be difficult to model and extract from the image, and they are usually housed in cluttered environments.

Most tracking methods are feature-based, that is, features of the object such as edges (Peng and Medioni 1988; Yang and Levine 1992) or corners (Roberts and Nashman 1992) that distinguish it from the background are extracted from the image and followed via the image sequence. Other methods are based on image differencing. They either segment moving objects from the background or construct an optical flow field. There are also methods based on correlations between pixels in successive images (Kokuer and Clark 1992).

In this application, the piglets are likely to be very much smaller than the total field of view of the camera. Therefore, optical flow methods are unlikely to be useful because they require an object being tracked to have a large, smoothly varying internal area, and do not work well at object boundaries. Methods relying on correlations between pixels, that is, matching a grey-level template of a piglet to the pixels in the next image, are likely to fail when piglets move through areas of different illumination in the pen. Image differencing between successive frames as a method of extracting the objects from the images is difficult to use because it cannot detect piglets that remain motionless for long periods. It finds only the moving edge of an object and responds to moving shadows associated with the objects.

Because piglets lack features such as corners, and their edges may contrast less with the background than many other edges in the scene, the only trackable feature is the outline of the piglet itself. This implies that the piglet images must be segmented from the background to be tracked.

Image differencing between the current image and an image of the background without the piglets does not suffer from the problems of differencing between successive frames and

Correspondence to: N.J.B. McFarlane

would provide a useful method of segmentation. However, in practice, it is difficult to obtain an image of the background without the piglets, since this would entail the removal of the animals from the pen. Therefore, such an image would have to be produced and the lighting changes updated with images in which the animals appear.

The object of this research was to find techniques for the image segmentation and tracking of piglets in a fairly simple environment, and to suggest how these might be further developed for tracking piglets in farrowing pens. A method was developed for producing and updating a time-averaged, median background image from images in which the piglets were present.

2 Description of image sequence

The images used in this research were simpler in several important ways than those obtained from a farrowing pen. The piglets were approximately 4 weeks old, and therefore had a better defined shape than newly born piglets. The floor on which they were moving was covered with plain straw, giving a simpler image background than the metal grid commonly used in farrowing pens. No sow was present, so that the piglets were the only moving objects in the image. The lighting conditions – diffuse overhead lighting – were better than the strong lights and hard shadows that might be found in a farrowing pen.

Figure 1 shows a typical image of the ten piglets present in the pen. Each piglet was marked by a stripe of dark ink along its back, but this feature was not clearly visible at the resolution used. The images were digitised from video tape with an initial resolution of 256×256 pixels, and the resolution was further reduced to 128×128 pixels by Gaussian blurring and shrinking. At this resolution, 128 pixels were equivalent to about 3.5 m on the pen floor. The test sequence of 200 images was captured and digitised at the rate of one image/140 ms – the fastest rate possible with the computing equipment available for this work.

Fig. 1. Typical image of piglets

3 Segmentation algorithm

The piglets did not possess sharp corners or unique markings; therefore, the only trackable feature of a piglet was the outline of the piglet itself. Thus, a segmentation of the image into piglet and background pixels was necessary to distinguish the low-contrast, often motionless, outlines of the piglets from the profusion of edges in the scene.

The first stage of image segmentation was image differencing. Each successive image was subtracted from a time-averaged reference image, and the difference image was thresholded. Piglet pixels were defined as those above the threshold; that is, the piglets were assumed to be significantly brighter than the background. Image differencing between the current frame and a reference image gave better segmentation results than differencing between successive frames because it did not produce false positives where a dark shadow had moved away from an area of background. It identified the whole area of the piglet, rather than just the leading edge, and it was able to locate piglets that were not currently moving.

The reference image was a running median of the image sequence produced by the following method. Each pixel in the reference image was incremented by one if the corresponding pixel in the current image was greater in value or decreased by one if the current image pixel was less in value. Each pixel in the reference image then converged to a value for which half the updating values were greater than and half were less than this value – that is, the median. This technique requires the storage of only one reference image, and is computationally inexpensive. The median was chosen in preference to the mean because of its better rejection of outliers in the distribution of pixel values. Hence, a piglet moving through part of the image would not change the median pixel values as rapidly as it would the mean. The median image, when finally converged, is an image of the background without the piglets.

One problem in constructing and updating the reference image was the tendency of the piglets to remain motionless for long periods of time. To make an initial reference image, it was necessary to use a sequence of images in which the piglets were all in motion. Motionless piglets were also a problem during the updating phase, since they were only slightly brighter than the background, and could be copied into the reference image in only about six updates. Reducing the frequency of updates would not eliminate this problem because the positions of the piglets were nonrandom; they would preferentially cluster around the feeder or the heat lamp, so that they tended to appear in the same places independently of the interval between updates. Therefore, each update of the reference image was assisted by a mask that covered the regions of interest around each located piglet and in which the reference image was not to be updated. This prevented the (located) piglets being written into the reference image, whilst retaining the ability of the reference image to adjust to changes in lighting levels.

The second stage of segmentation used a Laplacian operator to improve the separation of piglets that were clustered tightly together. The dividing lines between tightly packed piglets were often brighter than the background would have been in their absence, making it difficult for the image differencing to separate the individuals in such a group. The (unprocessed) piglet image was convolved with a Laplacian operator to distinguish between the negatively curved backs of the piglets, and the positively curved grey levels of the regions between them. The Laplacian operator was thresholded into regions of positive, zero and negative curvature, and the pixels initially labelled zero because they could not be assigned to a piglet or the background on the basis of curvature, were relabelled as positive or negative according to a majority vote of their nearest neighbours. The positively curved regions were then removed from the thresholded difference image. The resultant image yielded a segmentation of the piglet images from the background with improved separation of the individual piglets.

In using the Laplacian operator to detect objects with curved surfaces, it is important to match the scale of the operator to the scale of the object in question. If the width of the operator (in pixels) is much smaller than the object, it responds to noisy details on the object surface rather than the object itself; if it is larger, the object itself becomes unnecessarily blurred. In practice, the most convenient method of matching the scales of the object and the operator is to reduce the resolution of the image by blurring and shrinking without changing the size of the operator, since the hardware can perform the blurring, and the number of pixels to be processed is reduced. In this application, the image resolution of 128×128 pixels, as described in Sect. 2, was found to work well. At this resolution, the piglets were approximately 15 pixels long and 6 pixels wide.

The third stage of segmentation was to form the segmented pixels into meaningful groups. Chain coding was used to group the pixels into blobs. Blobs that were too small to be piglets were ignored as noise. Holes in blobs were also ignored to keep the representation simple. This could result in several partially joined piglets belonging to the same blob, which was not expected to cause problems because it was possible to wait for the piglets to separate before identifying them as individuals. Figure 2 shows the result of the segmentation as applied to Figure 1.

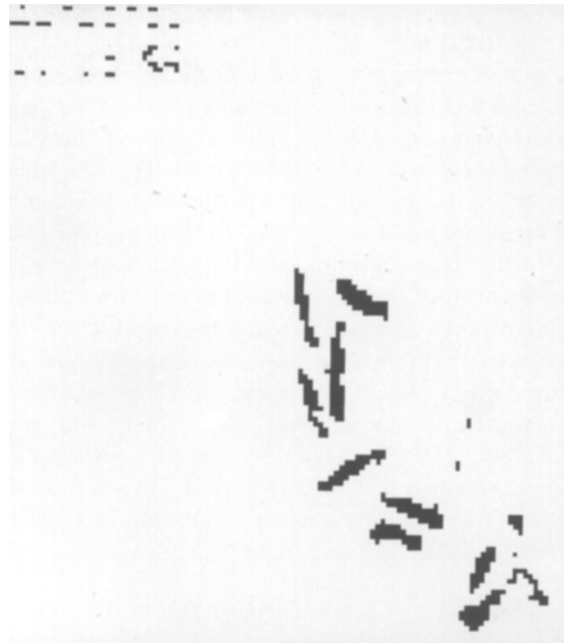


Fig. 2. Segmented image of piglets

4 The tracking algorithm

During tracking, the piglets were modelled as ellipses. A list of piglets was kept, listing for each the five ellipse parameters of x_c, y_c, a, b and α, where x_c and y_c were the coordinates of the centre, a and b were half the lengths of the major and minor axes, and α was the rotation angle. The list was initialised from the first segmented image. For each chain coded blob an ellipse was calculated from the means and covariances of the boundary coordinates. The centre coordinates x_c and y_c were given directly by the mean coordinates, and α was given by the direction of the principal eigenvector.

The parameters a and b were calculated by considering the eigenvalues of the covariance matrix that would be obtained from data based on the exact equation of an ellipse with length and width $2a$ and $2b$. In this case it can be shown that the eigenvalues λ_1 and λ_2 would be given by

$$\begin{aligned} \lambda_1 &= 2a^2 \\ \lambda_2 &= 2b^2 \end{aligned} \tag{1}$$

Hence, the parameters a and b were given by

$$\begin{aligned} a &= \sqrt{(2\lambda_1)} \\ b &= \sqrt{(2\lambda_2)} \end{aligned} \tag{2}$$

Since the eigenvalues of a covariance matrix were the variances of the data in the principal directions, it can be seen that a and b were proportional to the standard deviations of the coordinate data in these directions.

Having initialised the list of piglets from the first image, the primary objective of the tracking was to update the parameters of each piglet in the list according to the image sequence. One possibility for tracking was to use a hill-climbing technique to move each piglet ellipse into its updated position. However, this is computationally expensive because it requires several tests of the ellipse against the image pixels (either the grey-level pixels of the original image or the binary pixels of the segmented image) to determine the best new position. Another possibility was to compare the piglet ellipses with the new blobs at the level of the chain-code or ellipse representation, but this was difficult to do because the blobs did not necessarily correspond one-to-one with the piglets, despite the best efforts of the segmentation process. A single blob might comprise more than one piglet, or a piglet image might be fragmented into more than one blob.

The method used avoided these problems and was as follows: a search window was defined locally around the piglet ellipse, the window was scanned for blob edges, and the coordinates of the edgels were modelled as a new ellipse, for which we used the same mathematics described previously in initializing the piglet image. Thus, a new position for the piglet was calculated with only one comparison of the ellipse against the image, whilst avoiding the complexities of comparing ellipses with high-level blob features. The search window was a strip of pixels around the ellipse that was scanned radially from the centre of the ellipse at various angles. Its width was limited by the maximum number of pixels that a piglet could move between images. Thus, the tracking of a piglet comprised the scanning of the image within an elliptical window that was centred about the piglet's previous position and the updating of the piglet's parameters based on the blob edges found in the window.

If the blobs had always corresponded one-to-one with the piglets, this method would have been the same as a comparison of the piglets and the blobs, since the same edge coordinates would be found and the same mathematics applied. However, the method used was robust where the blobs were fragmented or joined because the limited size and shape of the window restricted the edges found to those that could reasonably belong to the piglet being tracked. At angles around the ellipse where the radial search failed to find an edge, the edge coordinates were filled in by the ellipse equation. That is, the edge was assumed to have remained in the same position. This prevented the calculation of the new ellipse being unbalanced by omitting the coordinates entirely. At this point, it was possible to assign a score to the piglet that was equal to the percentage of radial searches around the perimeter for which edges had actually been found. A low score indicated a poor match between the ellipse and the image, and suggested that the piglet being tracked had been lost.

Having obtained a measurement of the new parameter vector for the piglet, the piglet's parameters were updated. The parameters x_c, y_c and α were set directly to the new values since the measurement noise was smaller than the possible changes in these parameters during the 140 ms sampling interval. The parameters a and b, in which apparent changes were almost entirely due to measurement noise, were updated by only 30% of the difference between the measurements and the old values. A further refinement of the tracking algorithm was to centre the search window in the next image on a position extrapolated from the changes in x_c, y_c and α, assuming that the piglet would continue with its present velocity and angular velocity during the next sampling interval. This placed the search window in the most likely place to find a piglet moving with uniform velocity.

Given the ability to track individual piglets, the remaining task required of the tracking algorithm was to maintain the complete list of piglets. This part of the algorithm deals with situations such as identifying new piglets, losing track of piglets, and blobs splitting into more than one piglet image. Therefore, the tracking algorithm was required to make decisions regarding the addition and subtraction of piglets from the list and the splitting of piglet images.

The decisions were made by comparing the list of piglets with a list of candidate piglets formed by converting the blobs in the current image to ellipse representation. An ellipse was deleted from the piglet list if its score, defined as the percentage of its perimeter for which local edges had been found, was below 75%, indicating that the ellipse had lost touch with the piglet being tracked. If the a and b parameters of an ellipse exceeded predefined maxima, or if its area had fallen below a predefined size, it was also deleted. A candidate ellipse was added to the piglet list if its size was acceptable, and if it did not overlap the ellipses of any piglets already on the list. The splitting of piglet images into more than one individual was achieved as a result of the deletion and addition criteria without any additional computation. An ellipse containing two distinct piglets was first deleted from the piglet list, either due to the size or goodness-of-fit criteria, and the component piglets were added from the candidate list, since they no longer overlapped any existing ellipses. Figure 3 shows a flowchart of the segmentation and tracking as described in Sects. 3 and 4.

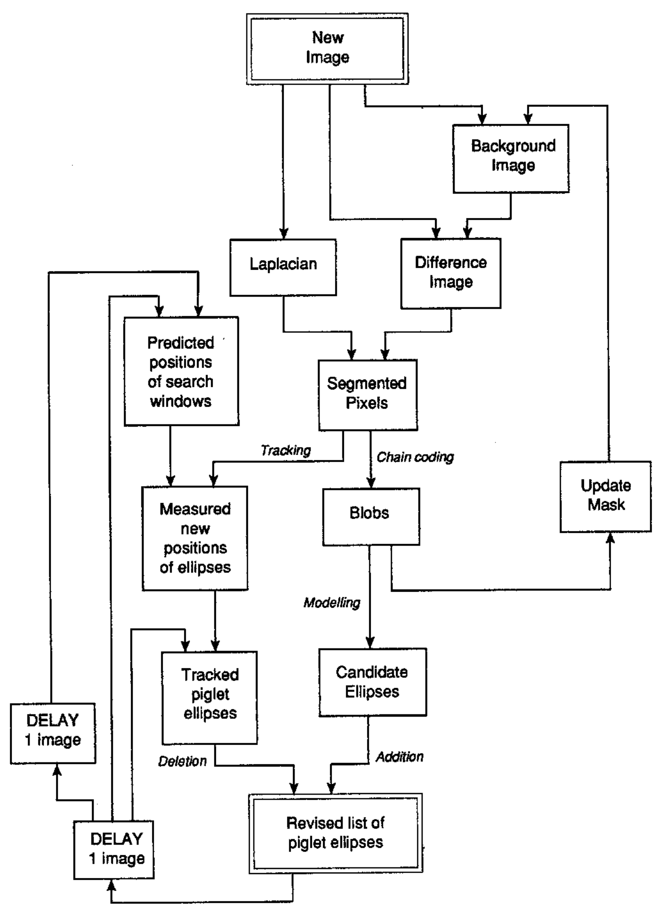
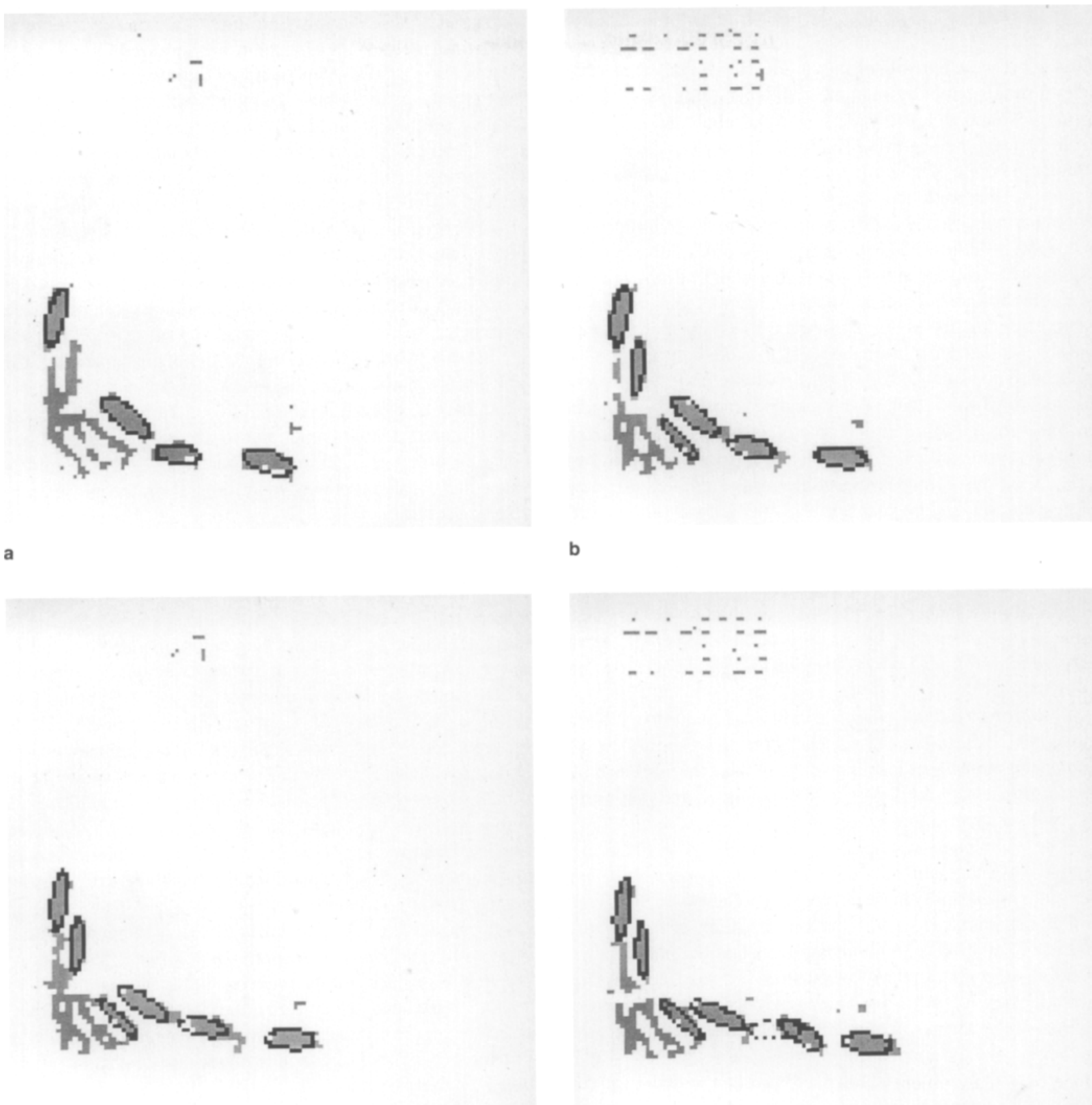


Fig. 3. Flowchart of the segmentation and tracking algorithm



a b c d

Fig. 4a–d. Tracking sequence

5 Results

The segmentation and tracking were tested on the sequence of 200 images described in Sect. 2. The piglets were tightly clustered at the start of the sequence and also between images 60 and 100. Between images 130 and 174, the piglets were well spaced out, but were almost motionless. During the rest of the sequence, at least some of the piglets were moving rapidly.

Over the entire image sequence, the average number of piglets being tracked at any one time was eight out of ten, and the mean number of images for which a piglet was tracked before being lost was eight. For images 60 to 100, in which the piglets were tightly grouped, these figures were respectively six out of ten and three. Figure 4 shows the segmentation and tracking applied to a short sequence, which illustrates some of the points of interest in this section.

A major point of difficulty with the image sequence was the speed of the piglets compared to the time between images; the piglets were capable of moving by nearly half their body length in the sampling interval of 140 ms – a speed of about $2\,ms^{-1}$. Worse still, the piglets could accelerate to full speed from zero in one sampling interval, making it impossible to predict the piglet's next position from its current position and velocity. This made the tracking difficult because the piglet could partially escape from the limited search window around the tracking ellipse. The tracking was successful when the piglets were moving slowly or moving with uniform velocity in a straight line, but contact with the nose of the piglet tendet to be lost when the piglet accelerated suddenly. Contact with the nose was also lost when the piglet rotated sharply or moved its head to one side, taking its head away from the long axis of the search window. Once the tracking ellipse had lost contact with the nose, it was left behind by the piglet until the match of the ellipse with the image data was so poor that it was deleted. At this point, the piglet blob was identified as a new piglet, and tracking began again. No attempt was made in this research to relate the new piglet image to the one that had just been deleted, but there were many cases in which such high-level information could have been used to help the tracking. Figure 4d shows where tracking has been lost in this way. The broken ellipse shows the last position of the second piglet from the right before tracking was lost due to its rapid acceleration and rotation.

It was observed that when tracking was lost due to a sudden movement of the piglet, the rear end of the piglet almost always remained within the ellipse. This suggests that the rear end may be the most easily trackable part of the piglet, and may form the basis of some future work.

The segmentation of the piglet images from the background was successful in that no false piglets were found, and there was no difficulty in identifying the piglets when they were widely separated. However, the segmentation of individuals from each other was sometimes poor when the piglets were close together, particularly when they formed large, tightly packed groups. This difficulty was partly due to the low resolution of the images, which sometimes made it difficult for the human eye to separate the individuals.

No problems were experienced with the updating of the background reference image, in that it continued throughout the sequence to represent the background grey levels reliably without any piglets being incorporated into the image. Initialising the reference image was more difficult, since it required a series of images in which the piglets were in random positions. The positions of the piglets in the 200 images were not random due to the 40 frames in which the piglets remained clustered together without moving, so a hand-picked and randomised sequence was used in the initialisation.

6 Further work

Future work on this tracking application centres on the need to improve the extraction of the features for tracking by better segmentation and the need to improve the tracking algorithm itself.

The only problems encountered in the segmentation process were in distinguishing the piglets from one another when they were tightly grouped. This was largely due to the reduced resolution of the images. The reduced resolution was important for blurring out spurious details on the surface of the piglets, such as the ink stripes noted in Sect. 2. However, the boundaries between the piglets were smaller-scale structures than the piglets, and could only be perceived reliably at the higher resolution. A better segmentation could probably be achieved by combining information from both resolutions in a multiscale approach.

Inevitably, some piglets are not distinguishable by pixel-based segmentation alone, and therefore, some knowledge of the shape of the piglet is required to recognise blobs that consist of more than one piglet and to separate blobs into their component piglets. This will require analysis of the boundary shape and local curvature. Hough-transform techniques may also be useful.

The major problem during tracking was the loss of tracking due to large, unpredictable movements of the piglets. The effect of these can be reduced if the image-capture rate can be increased to at least 1 image/50 ms to make it less likely that the piglet escapes from the tracking ellipse in a single sampling interval. Also, given a faster image-capture rate, the dynamics of the piglets would be more apparent, and they could then be modelled and tracked by state variable techniques such as the predictor-estimator or Kalman filter. If a faster capture rate is not possible due to the constraints of processing time, it should be possible to track by comparing the old ellipse positions with the new candidate ellipses, as well as with the low-level information of the segmented image. In most cases, this ought to result in an unambiguous match between ellipses in the candidate and piglet lists.

It was noted in Sect. 5 that the rear end of the piglet was almost always found inside the previous position of the tracking ellipse. This means that the rear end should be much easier to track than the head, which cannot be found reliably by a local search about the previous position, and suggests an algorithm that locates the rear end as a powerful first indicator of the piglet's new position.

Most of the techniques described in this research are applicable to the more complex scene of a farrowing pen. The age of the piglets and the floor pattern would probably make little difference to the performance, and methods could be developed to cope with difficult lighting conditions. The problems caused by the rapid motion of the piglets would be reduced in a farrowing pen because the piglets would be relatively slow-moving for the first hour or two after being born. The major problem in transferring this work to a farrowing pen is the presence of the sow. Techniques for distinguishing between piglets and parts of the sow are required, and the maintenance of a background image is more difficult with a large fraction of the image occupied by a moving sow.

7 Conclusions

The combination of image differencing with respect to a time-averaged background and convolution with a Laplacian operator was a good technique for distinguishing the piglets from the background.

The segmentation of individual piglets from one another was sometimes poor when the piglets were very close together, mainly due to the resolution of the images at 128 pixels/10 m being too low. Future work might use a higher resolution to distinguish the individual piglets, whilst retaining the benefits of reduced noise from the low resolution.

Tracking via a single scan of a search window around the piglet's predicted position was computationally inexpensive, but often became lost when the piglet's head moved outside the search window. This was mostly due to the large distances moved by the piglets during the 140 ms interval between images. Solutions to this problem might be increasing the image capture rate, using state-variable estimators to model the dynamics of the piglet, using higher-level information to aid the tracking, and using the easily located rear end of the piglet to improve the placement of the search window.

A technique was developed for maintaining an image of the median background over time, which requires the storage of only the reference image itself and a mask that prevents stationary piglet images being written into the background.

Image differencing with respect to a time-averaged background was better than differencing between successive frames because it did not produce false positives where a dark shadow had moved away from an area of background. It identified the whole area of the piglet, rather than just the leading edge, and it was able to locate piglets that were stationary.

In extending this work to the farrowing-pen environment, most of the techniques developed are probably applicable, but further techniques need to be developed to cope with difficult lighting conditions and the presence of the sow.

Acknowledgements. This work was funded by the Office of Science and Technology (OST) and the Ministry of Agriculture, Fisheries and Food (MAFF).

References

Deshazer JA, Moran P, Onyango CM, Randall JM, Schofield CP (1988) Imaging systems to improve stockmanship in pig production. Silsoe Research Institute, Div. Note DN1459, Wrest Park, Silsoe, Bedford, UK

Geers R, Goedseels V, Parduyns G, Nijns P, Wouters P, Bosschaerts L (1990) Integrated control of air and floor temperature in piglet houses: animal and engineering aspects. Ann Zootech 39:19–25

Kokuer M, Clark AF (1992) Feature and model tracking for model-based coding. International Conference on Image Processing and its Applications 354:135–138, 153

Mount LE (1968) The climatic physiology of the pig. Edward Arnold, London

Peng SL, Medioni G (1988) Spatio-temporal analysis of an image with occlusion. Image-understanding Workshop, Vols 1–2, 105, 433–442

Roberts K, Nashman M (1992) Real-time model-based tracking combining spatial and temporal features. J Intell Robotic Syst 5:25–38

Schofield CP (1993) Image analysis for non-intrusive weight and activity monitoring of live pigs. 4th International Symposium on Livestock Environment, University of Warwick, ASAE, 2950 Miles Rd, St Joseph, Michigan, USA, pp 503–510

Van der Stuyft E, Schofield CP, Randall JM, Wambacq P, Goedseels V (1991) Development and application of computer vision systems for use in livestock production. Comput Electronics Agriculture 6:243–265

Yang YH, Levine MD (1992) The background primal sketch: an approach for tracking moving objects. Machine Vision Appl 5:17–34

Nigel McFarlane received the degree of B.Sc. in Physics from the University of Birmingham in 1984 and the degree of M.Sc. in Modern Electronics from the University of Nottingham in 1986. In 1986 he joined Silsoe Research Institute, where he has carried out research in the fields of grain dryer control and image processing.

Patrick Schofield received the degree of B.Sc. in Agricultural Engineering from Silsoe College in 1977. In 1977 he joined Silsoe Research Institute, where his current research interests are in image processing applications for livestock production.