Motion Tracking of Cattle with a Constrained Deformable Model

Daisuke TSUTSUMI† Yasuyo KITA‡
† Department of Electronic Engineering, Hokkaido Industrial Research Institute
Kita 19, Nishi 11, Kita-Ku, Sapporo, Hokkaido, Japan 060-0819
‡ Research Institute of Intelligent Systems,
National Institute of Advanced Industrial Science and Technology (AIST)
AIST Tsukuba Central 2, 1-1-1 Umezono, Tsukuba, Japan 305-8568
tsutsumi@hokkaido-iri.go.jp, y.kita@aist.go.jp

Abstract

In this paper, we propose a method for monitoring the motion of cows by tracking the white patterns on them with constrained deformable models. As input for observation, we use image sequences of overhead views of cows taken by a camera installed on the ceiling of a breeding room. First, the 3D coordinates of the boundaries of the white patterns on the head, neck and torso parts respectively are registered. For each body part, the relationship between variations of the patterns in appearance and the parameters specifying the movement are defined in advance from that part's movement characteristics. Using deformable models which are constrained to be deformed based on these relations, the boundaries of the patterns are robustly tracked, producing as output the parameters specifying the movement of each part. Preliminary experiments using actual image sequences have shown the practical usefulness of the proposed method.

1 Introduction

Daily monitoring of the daily behavior of cattle is important in the detection of diseases and injuries in an early stage: for example, trembling may suggest the presence of BSE (Bovine Spongiform Encephalopathy). Since such symptoms are easy to overlook unless they are monitored over a long period, automatic monitoring systems which can watch and judge behavior instead of farmers are very desirable. Magee et al.[1] developed a method of tracking a cow's gait in some detail, using elaborate models for detecting lameness. Their method, however, assumes cows walk one at a time past a camera in a controlled fashion, and cannot be directly applied to the tracking of ordinary and arbitrary daily movements.

In this paper, we approach the problem for a more practical point of view, aiming to track the motion of cows in image sequences which are taken during their ordinary daily life. These imaging conditions present several dif - culties:

1. Cows often occlude each other
2. Foreground/background contrast is often poor
3. Cow motion is non-rigid

To avoid occlusion, we use top-down views taken by a camera installed on the ceiling of a breeding room as shown in Figure 5. Although leg movements cannot be observed in these images, they offer good visibility of the other main body parts, and observation of just these is thought to provide enough information to identify most abnormal motions. Regarding contrast, in Japan almost 90% of cattle is of the breed Holstein with white patterns against a black body. While the black body has low contrast against the typically dark ground which serves as a background in the top-down views, the white patterns can be robustly extracted. Therefore, we use the white patterns as tracking targets rather than the contour of a whole cow. One drawback of this strategy is inapplicability to cows without patterns, though the fraction of such cows in Japan seems less than 10 %. An advantage, however is that identi cation of individual cows can be tackled in an integrated fashion, since the patterns are unique to each individual.

We simplify the problem of non-rigidity by modelling a cow as three segments. Owing to the relatively simple appearance of the top view of a cow, the appearance can be divided into non-overlapping head, neck and torso parts. This separation enables to describe the movement of a cow simply. The head and torso parts can be treated as moving rigidly, while for the neck part a six degree-of-freedom rigid transformation and a bending transformation can represent its movement. How the appearance of the white patterns on each part is affected by this movement can be calculated, and therefore observation of these allows us to infer the movement.

Proceedings of the 16 th International Conference on Pattern Recognition (ICPR'02)
1051-4651/02 $17.00 © 2002 **IEEE**

We turn now to the extraction of the white patterns from images. For extracting the boundary of an arbitrary region from an observed image, deformable models such as snakes [2] are powerful tools. However, in the general case it is difficult to adjust the stiffness and smoothness parameters of a snake so that it fits a particular shape well. Active shape models [3] provided a solution to this problem by providing a framework for learning the possible variations in shape deformation from examples given in advance; from then on the deformable model is only deformed according to these variation rules and will efficiently be able to find similar shapes in new images. Although the learning process involved is a little troublesome, robustness against noise and occlusion is much improved.

In case of tracking a limited action such as “walking”, spatiotemporal models obtained from training data on typical motion work effectively both for robust tracking and for analysis of the action [4][1]. For tracking free actions, however, both training and usage of such spatiotemporal models become too troublesome and complicated to realize.

In our case, since the shape variation of the white patterns is caused by the movement of each part which can be described simply, it can be represented by a functional formula without learning from examples. Using simple movement models for each part, we calculate the relationship between variations of the white patterns in appearance and parameters specifying the movement and use this as a constraint in a deformable model for tracking the patterns. The white patterns on a part are robustly extracted by the deformable model, and at the same time the parameters specifying the movement of the part are obtained. The idea of our constrained deformable model is close to deformable templates [5] which deforms in a constrained shape-space. The differences are 1) our method deals with more general situations in a uniform manner and 2) the deformable model parameters directly represent the motion of a tracking object.

In Section 2, the principles of the monitoring system are described. Some preliminary experiments using image sequences are presented in Section 3.

2 Method of monitoring

2.1 Principles

Image sequences obtained from a camera mounted on the ceiling of a breeding room are shown in Figure 5. The motion of a cow is determined by tracking white patterns respectively for the head, neck and torso. Each part has its own coordinate frame as shown in Figure 1. Integration of the movements of the three parts enables recognition of high-level behaviors. However, in some cases even the individual movements of single parts can point to abnormality, for example if the cow is trembling.

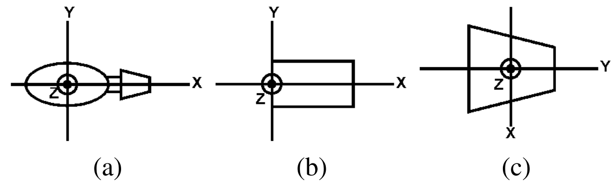


Figure 1. Coordinate frame of each body part: (a) Torso (b) Neck (c) Head

In order that each cow can be recognized, before tracking occurs the 3D coordinates of the boundaries of white patterns on each cow are obtained using the two views taken simultaneously from the top and the side, and are registered respectively for head, neck and torso parts. Although these processes are carried out manually at present, we are considering methods of automatic registration.

The monitoring process is divided into two modes: 1) Locating and 2) Tracking a cow. The first mode locates a cow in the scene, identifies it and outputs information on its initial position and posture. The second mode starts when the first mode has identified a target and performs continuous tracking. Running several processes in parallel allows more than one cow to be tracked at the same time.

2.2 Constrained deformable model

Suppose that a part whose movement can be specified by the parameter vector $\mathbf{q} = (q_1, q_2, ...q_m)^T$ is observed by a camera. When all other viewing parameters (including camera parameters) are known, the observed image coordinates corresponding to n points of the part, $\mathbf{u} = (u_1, v_1, u_2, v_2, ...u_n, v_n)^T$ can be expressed as a function of $\mathbf{q}$:

$$\mathbf{u} = \mathbf{F}(\mathbf{q}) \tag{1}$$

We assume that $\mathbf{u_o}$ is observed, $\mathbf{q_c}$ is an initial (close) estimation $\mathbf{q}$ of its state, and $\mathbf{u_c}$ contains the projected coordinates at $\mathbf{q_c}$ as shown in Figure 2. By expanding Equation 1 in a Taylor series around $\mathbf{u_c}$ and taking terms up to first order, we obtain:

$$\mathbf{u_o} = \mathbf{u_c} + \left.\frac{\partial \mathbf{F}}{\partial \mathbf{q}}\right|_{\mathbf{u_c}} \Delta\mathbf{q} \; . \tag{2}$$

Then, the $\mathbf{q}$ consistent with $\mathbf{u_o}$ is obtained by adding the following $\Delta\mathbf{q}$ to $\mathbf{q_c}$:

$$\Delta\mathbf{q} = \left[\left.\frac{\partial \mathbf{F}}{\partial \mathbf{q}}\right|_{\mathbf{u_c}} \right]^{\dagger} (\mathbf{u_o} - \mathbf{u_c}) \; . \tag{3}$$

Proceedings of the 16 th International Conference on Pattern Recognition (ICPR'02)

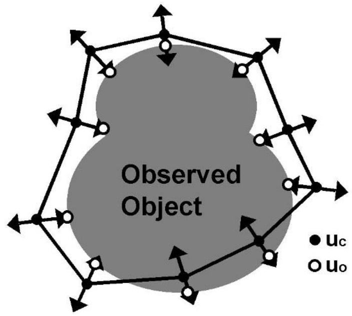


Figure 2. Search for observed edge points

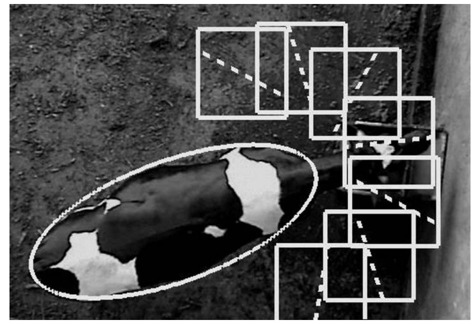


Figure 3. Search for head part

Here, $[\mathbf{A}]^{\dagger}$ represents the pseudo-inverse of matrix $\mathbf{A}$. Because of the rst-order approximation leading to Equation 2, the obtained solution may include some errors, and an iterative calculation using the obtained $\mathbf{q}$ as a new $\mathbf{q_c}$ for each step leads to convergence to the correct state.

We can see that $\mathbf{u_c}$ is deformed while maintaining consistency with the constrained part movement, and so we adopt $\mathbf{u_c}$ as our constrained deformable model. Looking in more detail, suppose that at a given time its shape is given by $\mathbf{u_c^t}$ derived from model con guration $\mathbf{q_c^t}$:

1. $\mathbf{u_o}$ is determined as the strongest edge point in the normal direction of $\mathbf{u_c^t}$ within a xed distance as shown in Figure 2.
2. $\mathbf{q_c^{t+1}}$ and $\mathbf{u_c^{t+1}}$ are calculated from the value of $\Delta\mathbf{q}$ obtained by putting $\mathbf{u_o}$ and $\mathbf{u_c^t}$ into Equation 3.

These two steps are iterated until convergence. The algorithm is similar to that used in active shape models [3], except in the method of constraining the shape of the deformable contour. As the contour de ned by $\mathbf{u_c}$ deforms, the change in $\mathbf{q_c}$ represents the movement of the part.

2.3 Locating mode

When searching for a cow in an image, the white patterns on its torso are searched for rst since they are relatively easy to nd because of their larger areas and slower movement than the head and neck. Since cows are usually standing or walking, their torsos are parallel to the ground most of the time, and the appearances of the white patterns viewed from the z direction (refer to Figure 1(a)) are used as 2D templates for the search. Once a large white region is found in the image, it is compared with the known 2D templates to identify the cow and to determine the position and posture, here the direction in the horizontal plane, of its torso.

Then, the white patterns on the cow's neck and head are searched for in the regions within which the deformation model permits them to lie. Figure 3 shows an example of a search for head patterns. Once a white region is found in one of the local areas (the white squares in Figure 3), the typical appearance of the white pattern is overlaid on the region. Here, by typical appearance we mean the appearance of the white pattern when the cow is standing in its most common pose, and its direction in the image plane is set according to the head direction predicted from the relation with the torso which is shown by dashed lines in Figure 3. Using the typical appearance to initialise the constrained deformable model, the contour of the observed white region is extracted. The resultant shape becomes the initial state for the tracking mode.

2.4 Tracking mode

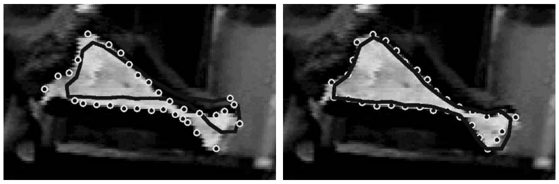
Once the initial state of a cow is found, the white patterns are tracked in the following frames by using the state of the previous frame to initialise the constrained deformable model. Note that one unit of the model corresponds to white patterns on a body part rather than a single connected white region.

Although a torso part is searched for with the constraint that it is parallel to the ground in “Finding mode”, any rigid transformation can be taken into consideration in “Tracking mode”. The movement of a cow's head is also represented by a rigid transformation. Therefore, its parameter vector is $\mathbf{q} = (t_x, t_y, t_z, r_x, r_y, r_z)^T$, where t_x, t_y, t_z and r_x, r_y, r_z represent translation along the (x, y, z) axis and rotation around the (x, y, z) axis respectively. For the neck, we plan to add bending parameters, though this has not yet been implemented.

3 Experiments

We applied the proposed method to image streams monitoring the corner of a breeding room containing a water trough. Since cows come to drinks one by one, this is one of the best places for watching them. In this experiment, the motions of two cows having white patterns only

Proceedings of the 16 th International Conference on Pattern Recognition (ICPR'02)
1051-4651/02 $17.00 © 2002 **IEEE**



(a) (b)

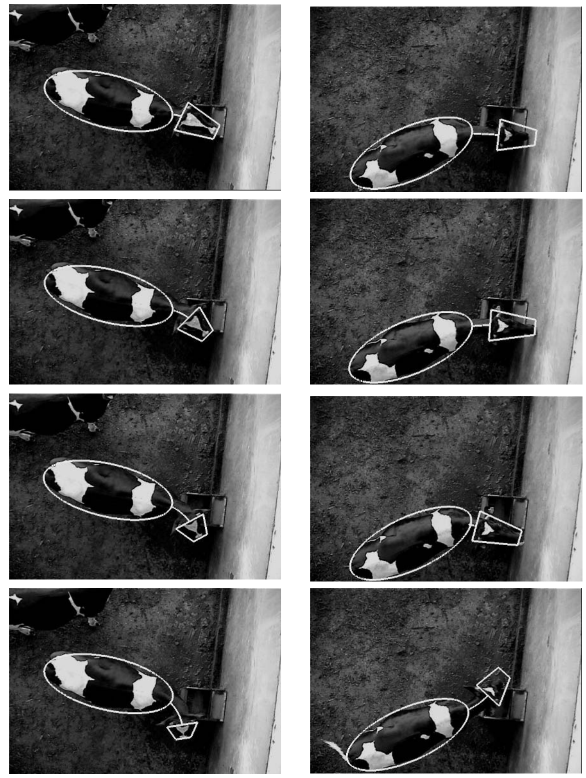
Figure 4. Extraction of white pattern by constrained deformable model: (a) initial state (black line) and detected edge points (white points); (b) final state after convergence (black line).

on their heads and torsos were tracked; as a result, no attempt was made to track their necks. The image sequences are good examples for testing the tracking capabilities, because large movements of the head while drinking are included.

Figure 4 shows an example of the process of finding the initial state for “Tracking mode”. At this stage, the cow had already been identified and the position and direction of its torso had been determined. The black line in Figure 4(a) shows the initial state of the constrained deformable model obtained from the typical appearance of the head's white pattern. The white points with a black dot in its center in Figure 4(a) shows the edge points obtained by the process explained in Figure 2. Figure 4(b) shows the converged state of the deformable model after 50 iterations. As shown here, the constrained deformable model robustly extracts the white pattern and the initial state of the head: (33.7, 2.5, 26.0) degree rotation around the (x,y,z) axes.

Figure 5 shows two examples of tracking results. The position and posture of the torso are represented by a ellipse in the image plane, since the bodies are always almost parallel to the floor during monitoring. The movement of the head part is represented by a trapezoid which is superimposed on the front of the cows' faces. Currently, a line connecting the two parts is drawn based on a simple rule and does not represent the neck movement accurately. Note that the height of the head relative to the torso cannot be estimated at present, although we are studying ways of estimating this using the relationships between parts such as the neck length.

Change in the estimated angles of the head during tracking (cow B in Figure 5) are shown in Figure 6. In some parts of the image sequences, the correct angles around the x and the z axes (refer to Figure 1(c)) can be manually measured from a top view and the side view taken simultaneously. In Figure 7, the resultant angles during such periods are plotted as well as the manually measured angles; the angles are shown with “+”,“×” and “•”, “◦” respectively for the x and the z axes; the solid and dashed lines



(a) (b)

Figure 5. Results of tracking: (a) cow A; (b) cow B.

illustrate the change in the estimated and measured angles respectively.

Figure 8 shows the relation between the estimated errors and the rotation angles; the errors are shown with “+”,“×”,“•” and “◦” in the same manner as in Figure 7. A very large rotation around the x axis makes the appearance of the white patterns so small that the tracking is impossible. As a result, tracking rotation around the x axis is limited to about 70 degrees. At present, our method gives up tracking when the area of the white patterns becomes smaller than a fixed threshold, and moves into “Locating mode”. The larger errors seen in the term of 0∼30 degrees of the rotation around the x axis come from the fact that the appearance changes quite slowly during this term. However, even in this case, the errors are at most 15 degrees. In total, the average errors are 6.9 degrees in rotation around the x axis and 5.5 degrees around the z axis.

As shown in Figure 7, the estimated errors stably change with tracking time so as not to cause spurious jag on the resultant angle curves. In other words, the estimated curves are quite similar to the measured ones aside from bias errors. We think this characteristic is very important for the purpose of judging behavior of cows.

Proceedings of the 16 th International Conference on Pattern Recognition (ICPR’02)
1051-4651/02 $17.00 © 2002 IEEE

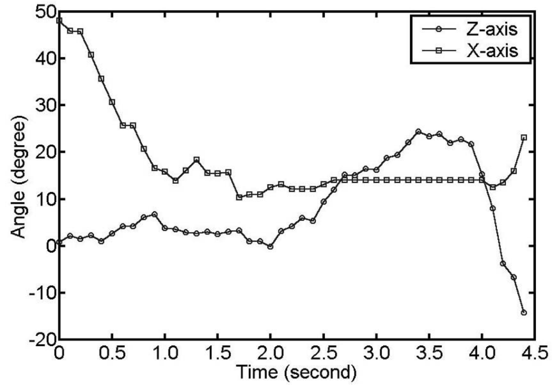


Figure 6. Change of angles during tracking (cow B)

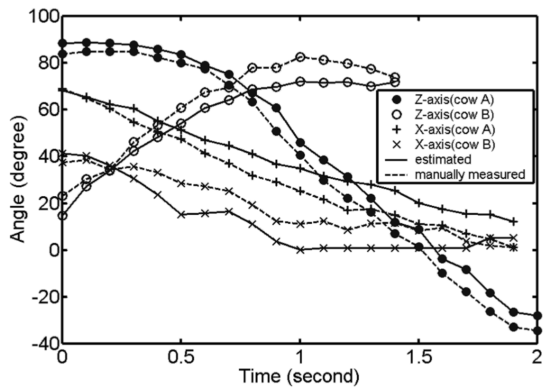


Figure 7. Change of estimated and manually measured angles during tracking

4 Conclusion

In this paper we have proposed a method for monitoring the daily movement of cattle by tracking the white patterns on them with constrained deformable models. By using top-down views, the torso, neck and head parts of a cow can be clearly observed without occlusions. The relationships between the appearances of the white patterns on each part and the movement of the parts are derived and used to robustly track the patterns while determining the movement.

In the experiments shown in this paper, large movements of a cow's head were successfully tracked. The errors in the estimation of the head posture were generally less than 10 degrees. Apart from bias errors in the absolute posture, the estimated motions are very close to the correct (manually-measured) ones. This shows the proposed method has a bright prospect for realizing an automatic monitoring system which can watch and judge cattle behavior. In future work, we will make extensions for dealing with more complicated movements like bending of the neck. Other aspects we will consider include integration of the movements of each part and assembling a practical monitoring system.

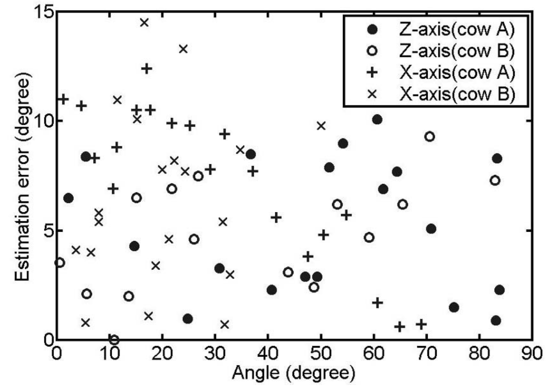


Figure 8. Error in estimated angles of head

Acknowledgments

The authors wish to thank Dr. Kazuo Tanie, Dr. Katsuhiko Sakaue and Dr. Nobuyuki Kita for their useful advice, Mr. Takeshi Hori and Dr. Keigo Kuchida for their help in providing us with the sequential images, Dr. Andrew J. Davison for his comments to improve English quality, and the reviewers for their helpful comments. This work was studied at the National Institute of Advanced Industrial Science and Technology (AIST) where the rst author stayed under its technical training program.

References

[1] D. Magee and R. Boyle: "Feature tracking in real world scenes (or how to track a cow)", In *IEE Colloquium on Motion Analysis and Tracking*, pp. 2/1–2/7, 1999.

[2] M. Kass, A. Witkin, and D. Terzopoulos: "Snakes: active contour models", *International Journal of Computer Vision*, Vol. 1, No. 4, pp. 321–331, 1988.

[3] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham: "Active Shape Models –Their Training and Application", *Computer Vision and Image Understanding*, Vol 61, No. 1, pp. 38–59, 1995.

[4] A. Baumberg and D. Hogg: "Generating spatiotemporal models from examples", *Image and Vision Computing*, Vol. 14,no.8, pp. 525–532, 1996.

[5] A. Blake and M. Isard: *"Active contours"*, Springer-Verlag, London, 1998.