

Detecting Lameness in Livestock Using 'Re-sampling Condensation' and 'Multi-stream Cyclic Hidden Markov Models'

Derek R. Magee and Roger D. Boyle
School of Computer Studies
University of Leeds
Leeds, LS1 9JT,UK
drm@scs.leeds.ac.uk

Abstract

A system for the tracking and classification of livestock movements is presented. The combined 'tracker-classifier' scheme is based on a variant of Isard and Blakes 'Condensation' algorithm [6] known as 'Re-sampling Condensation' in which a second set of samples is taken from each image in the input sequence based on the results of the initial Condensation sampling. This is analogous to a single iteration of a genetic algorithm and serves to incorporate image information in sample location.

Re-sampling Condensation relies on the variation within the spatial (shape) model being separated into pseudo-independent components (analogous to genes). In the system a hierarchical spatial model based on a variant of the Point Distribution Model [16] is used to model shape variation accurately. Results are presented that show this algorithm gives improved tracking performance, with no computational overhead, over Condensation alone.

Separate Cyclic Hidden Markov Models are used to model 'Healthy' and 'Lame' movements within the Condensation framework in a competitive manner such that the model best representing the data will be propagated through the image sequence.

1 Introduction

In recent years there has been much interest in object tracking [6, 1, 16, 15, 12] and temporal modelling [8, 14, 18, 4]. The combination of object tracking and temporal modelling gives rise to many exciting application possibilities, for example; Isard and Blake [6] use a temporal model to improve the speed and robustness of their object tracker. Wren and Pentland [18] and Davis and Bobick [4] use temporal models to classify observed human movements and in the case of the latter use this information to trigger interactive responses in a virtual environment. Johnson *et. al.* [8] build a joint behaviour model in which a virtual human reacts in a realistic manner to observed behaviour in a limited domain. Sumpter and Bulpitt [14] use object tracking and temporal modelling to predict the behaviour of a flock of ducks or sheep for use in the control system of a robotic sheepdog.

We combine a stochastic ‘Re-sampling Condensation’ tracker with multiple ‘Cyclic Hidden Markov Models’ to model and classify livestock ‘behaviour’. In this paper we take lameness as an example ‘behaviour’ although it is hoped that this scheme may be extended to more complex animal behaviours such as the abnormal behaviour exhibited by animals when on heat (oestrus).

2 Background

2.1 The Condensation Algorithm

Isard and Blake’s Condensation (CONDitional DENSity propaGATION) algorithm [6] is a simple, yet powerful, way of propagating multiple dynamic problem solution hypotheses over time using a finite set of discrete samples. The original application of the Condensation algorithm was object tracking in an image sequence in which a model of an object (spatial and temporal) was fitted to an image sequence. Propagating multiple hypotheses gave improved robustness over single hypothesis techniques such as Kalman filters [3]. The basic Condensation algorithm is given below;

Repeat (for each sample at each time step):

- Select:** A sample is selected stochastically from the previous generation (time step) based on a ‘fitness’ measure.
- Predict:** A single state of this sample at the next time step is predicted using a stochastic temporal model.
- Evaluate:** The ‘fitness’ of the predicted sample location is evaluated.

Note: Samples may be initialised randomly or according to some known probability distribution for the first frame.

Condensation has also been applied to the field of temporal modelling (prediction and classification) of objects. Johnson and Hogg [9] use condensation to propagate multiple prediction hypothesis for pedestrian trajectory classification. Black and Jepson [2] use a similar scheme with multiple temporal models as a combined tracker and classifier to analyse an augmented whiteboard. Walter *et. al.* [17] use Black and Jepson’s method with continuous (non-discrete) Hidden Markov Models to classify the trajectories of people in an office scene.

2.2 Hidden Markov Models

Hidden Markov Models (HMMs) [13] are used to model observations from a stochastic process where there is some underlying structure, but observations are not deterministic. In many cases the exact nature of this process is not observable (i.e. ‘Hidden’), for example speech, however the resulting observations (sound in the case of speech) are. HMMs model the underlying process using a first order Markov chain and the relationship between the process and the observations by a probability distribution (either discrete or continuous). Figure 2.1 illustrates this by modelling the structure of the word ‘hello’ using one state for each syllable.

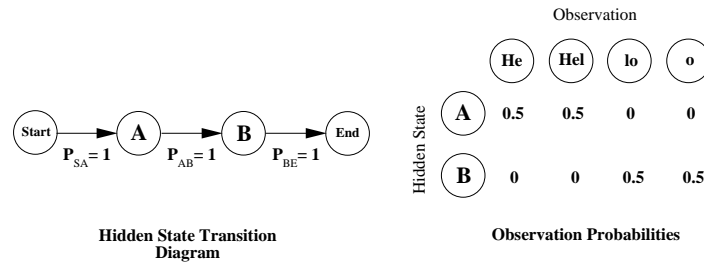


Figure 2.1: Modelling the word 'Hello' using a Hidden Markov Model

This model of the word 'hello' illustrates two important features of the HMM. Firstly, although the Markov chain is only first order, a temporal history is encoded by limiting the state transitions possible. In this example state A will always occur before state B thus encoding the temporal ordering of the two syllables in the word. Secondly, if the four possible permutations of syllables are examined ('He-lo', 'He-o', 'Hel-lo' and 'Hel-o'), it can be seen that an invalid combination is possible ('He-o'). If the HMM were to be used for recognition this would not be a problem as there is no other word in the English language that consists of these syllables, however if the HMM were to be used for speech generation this would present a problem. HMMs intended for sequence generation or prediction in general require more complex state transition architectures than HMMs intended for recognition only.

Rabiner in his excellent introduction to HMMs [13] lists three problems that need to be solved in order to use HMMs. Problem 1 is the evaluation problem, or how the probability that an observed sequence was produced by a given HMM may be calculated. Problem 2 is the recovery of a 'correct' sequence of hidden states for an observation sequence. Problem 3 is the training problem, or how to optimise a set of HMM parameters so as to best describe how an observation sequence (or sequences) come about. In addition to these there is a fourth problem; prediction. The problem is, given some knowledge of the system in the past / present, how can we predict the future behaviour of the system. This is an important, but not widely covered, problem. This is possibly due to speech recognition (the principal original use of HMMs) requiring no prediction from the HMM. Prediction is essential to our application and is discussed in section 4.

3 Building Spatial Models for Tracking Applications

Previous work has described how multiple contour models of an object class may be built [10] and how the variation within these models can be separated into independent components [11]. These models are a variation on the Point Distribution Model [16] in which a contour is modeled by a mean shape (described by a set of points) and a set of linear 'modes of variation'. We have applied this scheme to modelling livestock, building multiple models of cow outlines and separating the variation of each model into inter-animal, front legs and rear legs components as shown in figure 3.1.

Each model is discretised by performing vector quantisation on the projections of the training data in the eigenspaces of the model. This results in a set of prototypical 'states'

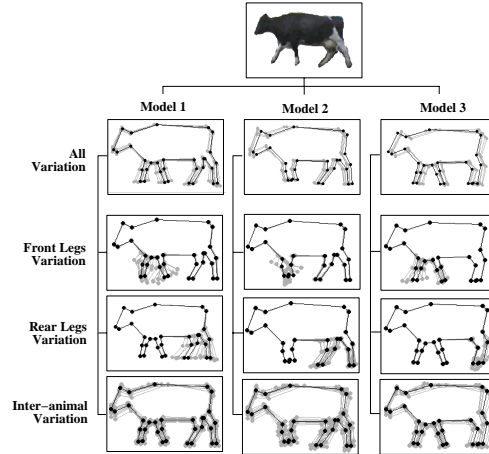


Figure 3.1: Separating Object Variation Using a Hierarchical Scheme

for inter-animal, front legs and rear legs components.

4 Building Multi-stream Cyclic Hidden Markov Models

Cyclic Hidden Markov Models (CHMMs) use a hidden state architecture in which the first and last states are joined. This is illustrated in figure 4.1.

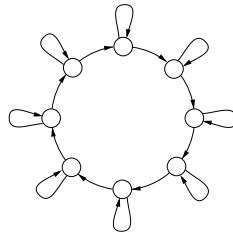


Figure 4.1: CHMM Hidden State Architecture

The Multi-stream Cyclic Hidden Markov Model (MSCHMM) is a CHMM with multiple sets of (discrete) visible states which model observationally independent but related features of a system. In our example we model the front and rear leg pairs as two separate ‘observational streams’ based on a single underlying CHMM. It should be noted that continuous (e.g. Gaussian mixture model based) observation probability distributions could be used, however these would be computationally more expensive to evaluate and as such are not used in the system implemented.

Hidden state transition probabilities are initialised by defining parameters P_s , P_c and P_x as the probability that the next hidden state remains the same, changes or the current state is the last state in the sequence respectively. If we assume initially that these param-

ters are the same for each hidden state we get the probability distribution given in equation 1 for the cycle length in states.

$$P(n) = P_s^r P_x P_c^N \times^{(N+r-1)} C_r = P_s^N P_x P_c^r \times \frac{(N+r-1)!}{(N-1)!r!} \quad (1)$$

Where:

- n = The Sequence Length
- $P(n)$ = The probability of a sequence of length n resulting from the CHMM
- N = The No. of Hidden States in the CHMM
- r = The No. of Hidden State Repetitions (n-N)

Training sequences are parsed into cycles (of various lengths) based on spatial model transitions. The number of hidden states (N) is selected to be the minimum cycle length in the training set. Values of N below this value may be used giving a CHMM with more generality and less specificity. Values of N above this value are not used as this would not allow generation of all training sequence lengths. An initial estimate of P_x as the reciprocal of the average sequence length is used and given $P_c + P_s + P_x = 1$ an exhaustive one dimensional search is performed in P_c (with fine quantisation) to minimise the square difference between the theoretical probability probability distribution of cycle length ($P(n)$) and a distribution calculated from the training sequences.

Initial visible state probability distributions are estimated by aligning each training sequence cycle with the set of hidden states, ‘time stretching’ such that the training sequence length is equal to the number of hidden states in the CHMM. Probabilities are estimated from the relative number of observations lying completely or partially over each hidden state as shown in figure 4.2.

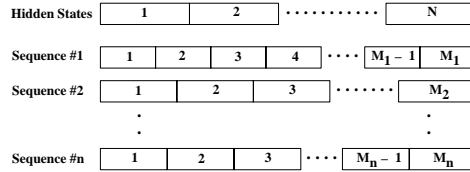


Figure 4.2: Initial Estimation of Observation Probabilities

The transition and observation probability estimates are improved using the Baum-Welch re-estimation method [13].

5 Object Tracking Using ‘Re-sampling Condensation’

The main drawback with the Condensation algorithm is that sample location is determined purely by prediction from past observations. Isard and Blake explain [7] that this results in sample locations clustering round regions of predicted high probability with few samples representing areas of lower predicted probability. This is a problem in applications where there are multiple possible outcomes of differing probabilities. Other sampling algorithms such as Markov Chain Monte Carlo [5] and Genetic Algorithms do not exhibit this problem

as they iteratively re-sample the solution based on previous results. These algorithms are however unsuitable for ‘real time’ applications due to their high computational cost.

‘Re-sampling Condensation’ is a novel two stage algorithm with no computational overhead over the standard Condensation algorithm that, under certain circumstances, gives more robust tracking results as image information is included in sample location. Re-sampling Condensation splits the samples into two groups. Sampling of the first group is performed using standard Condensation (as in section 2.1). Multiple samples are then selected stochastically from this initial sample set (based on a fitness function) and combined to give a new sample location. This is analogous to a single iteration of a genetic algorithm. In the livestock example three of the initial samples are selected and inter-animal, front legs and rear-legs characteristics taken from different samples to make up the new sample. Position and scale are determined stochastically from the mean and standard deviation of these parameters in the initial samples selected.

It should be noted that Re-sampling Condensation relies on the variation within the spatial model being separated into pseudo-independent components (position, scale, inter-animal, front legs and rear-legs characteristics in the livestock example). These are analogous to the ‘genes’ in a genetic algorithm. Experimental results (see section 7) suggest that the more components that a spatial model may be separated into, the better the final solution. This is intuitive as, considering a system with 2 independent parameters with N and M possible states for each parameter, searching using a model with a single parameter results in a search space of size $N \cdot M$ where as searching using a model with two parameters results in two search spaces of size N and M respectively. In most tracking examples parameters are not truly independent, however component separation can still yield improved search efficiency.

6 Combined Tracking and Behaviour Analysis using Multiple CHMMs

In section 2.1 it was described how the Condensation algorithm has wider application than simple object tracking. In particular Black and Jepson [2] use Condensation to track and classify the trajectories of a coloured whiteboard marker. In their scheme multiple trajectory models are used and the ‘gesture’ classified as one of six actions to be performed.

In our scheme we use the combined tracker and classifier paradigm of Black and Jepson to model object shape change over time in a ‘Re-sampling Condensation’ framework (see previous section). In our example ‘normal’ and ‘lame’ behaviours are modelled by separate CHMMs (see section 4) within the ‘Re-sampling Condensation’ framework. Samples are allocated to each model initially in even proportion and the CHMM is propagated through time with the sample. Over time the CHMM that best fits the observed object ‘behaviour’ will dominate (i.e more samples will be associated with that model). A simple classification can then be performed by comparing the number of samples associated with each model. A more complex classification method involves summing the posterior probabilities (relative fitnesses) for samples associated with each CHMM.

7 Evaluation and Discussion

7.1 Evaluation of Tracking Results

‘Re-sampling Condensation’ was evaluated using a test set of 10 sequences of healthy cows walking from right to left in a farmyard setting as shown in figure 7.1. These sequences are of approximately 5 seconds in length at 25fps and contain at least three complete cycles of the CHMM. Ground truth about these sequences was obtained by hand fitting landmark points. Other sequences were also used in the construction of the spatial (shape) model.



Figure 7.1: Typical Livestock Tracking Scenario

A set of leave one out tests was performed by building a CHMM from 9 of these sequences and using the remaining sequence as a test sequence. The training data was obtained by projecting the hand fitted ‘ground truth’ points into the model parameter spaces and selecting the nearest vector quantisation prototype (a crude but reasonably effective method). Tracking was performed at various levels of re-sampling (i.e. number of re-samples vs. number of condensation samples) from 0% (Normal Condensation) to 95%. The tracking results for the sample of maximum fitness were compared to the hand fitted ground truth, and statistics such as mean error gathered. Some results at different numbers (250, 500 and 1000) of total samples are given in figure 7.2.

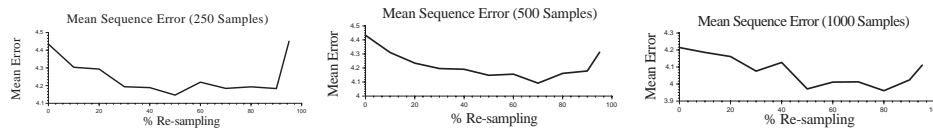


Figure 7.2: Mean Error (pixels) of Tracker at Different Levels of Re-sampling

The results in figure 2 show lower average error rates for Re-sampling Condensation at intermediate levels of re-sampling, although there is a drop off in performance at high levels of re-sampling. It should be noted that the tracking error is superimposed on the quantisation error inherent in the quantised model. The average error for the quantised ground truth data used for training is 4.08 pixels using the crude quantisation method described previously. It is encouraging that the tracker can, at optimal re-sampling, improve on this value. Error standard deviation shows similar trends to mean error although this is always higher than for the quantised training data. Examining individual sequences around the optimum operating points indicated by the graphs in figure 7.2, the Re-sampling Condensation tracker performs better than Condensation alone (0% re-sampling) in 80-100%

of cases. Even at very low (<20%) and very high (>90%) levels of re-sampling the Re-sampling Condensation tracker can perform better in more than 50% of cases.

These results are as would be expected as the inclusion of current image information in sample location improves tracking performance. The fall off in performance at high levels of re-sampling is due a ‘gene deficiency’ in the initial Condensation set of samples when few samples are allocated to the Condensation stage. It may be the case that this initial ‘population’ of samples contains no members with the ‘most correct’ individual characteristics and as such the most correct (highest fitness) solution cannot be found by the re-sampling stage. It may also be the case that the initial population of samples contains a high proportion of members with the most correct individual characteristics, in which case tracking performance will be particularly good. It is as such not desirable to operate at high levels of re-sampling as tracking robustness is lower than at intermediate levels of re-sampling. It should also be noted that the optimum level of re-sampling increases with the total number of samples due to this phenomenon.

7.2 Evaluation of Lameness Classification

The evaluation of the scheme in the detection of lameness was performed on sequences of humans walking as insufficient amounts of lame cow data were available. The 11 (healthy) subjects were asked to perform a choreographed lame walking motion in addition to their regular walking motion. Two sets of each were taken for each person and a set of leave one out tests performed. The spatial (shape) model used was a B-spline based model (*a la* Baumberg and Hogg [1]) rather than the straight line approximation as used for the cows as people are not well approximated by straight lines. The sum of the posterior probability (normalised fitness) for each CHMM was recorded over time for each sequence. Averaging the probabilities over time (excluding the first cycle to allow for initialisation of the tracker) gives an indication of the relative probability that each sequence is either ‘Healthy’ or ‘Lame’.

Using this method all 44 sequences were correctly classified and in all but two cases the results were very clear as can be seen in figures 7.3 and 7.4.

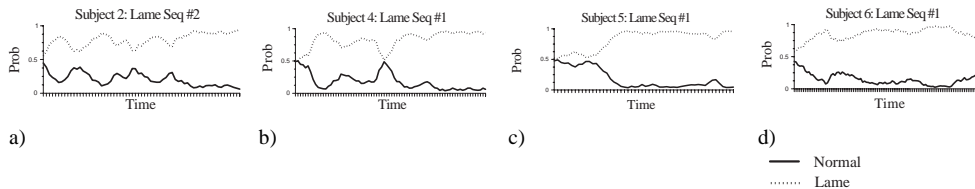


Figure 7.3: Model Probability vs. Time for ‘typical’ Lame Walk Sequences

Two sequences produced results which, although they were classified correctly, were not as clear (see figure 7.5). On investigation this was found to be related to the ‘lame’ portion¹ of the ‘lame’ CHMM being similar to the healthy walking motion in the rotationally normalised shape representation. The spatial model used to model the cows is not normalised by rotation and thus would not be subject to this problem. In the non-typical

¹ N.B. The ‘lame’ walking cycle consists of approximately half a cycle that is identical to a ‘healthy’ walk and half a cycle that differs. In the ‘healthy’ walking cycle the two half cycles are identical from the view used.

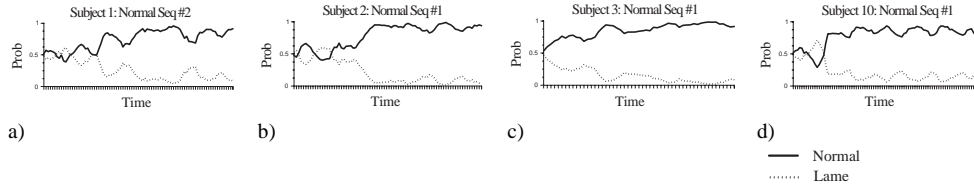


Figure 7.4: Model Probability vs. Time for ‘typical’ Normal Walk Sequences

lame sequence a significant number of samples were being propagated half a cycle out of phase with the actual cycle resulting in poor tracking performance. Similarly for the healthy sequence samples relating to the ‘lame’ CHMM were being propagated in both phases resulting in reasonable tracking by the lame model compared to the healthy model. These problems would be eliminated by using a non-rotationally normalised model or including relative rotation as an additional stream in the CHMM as out of phase samples would not be propagated.

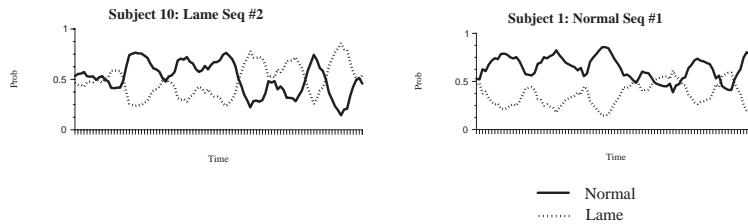


Figure 7.5: Model Probability vs. Time for the two ‘non-typical’ Sequences

8 Conclusions and Further Work

We have presented an extension to Isard and Blake’s Condensation algorithm that includes current image information in sample location by splitting samples into two sets, a ‘Condensation set’ and a ‘Re-sampling set’. The re-sampling stage relies on variation within the spatial model used being separated into pseudo-independent components which are then combined stochastically based on the Condensation set ‘fitness’ in a manner similar to a single iteration of a genetic algorithm. Experimental results show improved tracking performance over Condensation alone with the existence of an optimal proportion of re-sampling. We have also shown how multiple Cyclic Hidden Markov Models (CHMMs) may be used to model the temporal changes in object shape that represent a walking ‘behaviour’. These multiple CHMMs may be included in the Re-sampling Condensation framework to form a combined tracker and behaviour classifier. Experimental results using this scheme to classify the difference between a normal human walking motion and an artificial lame motion for unseen persons are very encouraging although insufficient ‘lame’ data was available to evaluate this scheme on Livestock. This combined tracker and classifier paradigm has very wide applicational scope beyond livestock monitoring and could be applied to areas such as gesture and gait recognition in future work.

9 Acknowledgments

This work was funded by the BBSRC. The authors would like to thank Prof. M. Forbes from the Leeds University dept. of Animal Physiology and Nutrition and the staff at ADAS Bridgets research institute for their help with data collection.

References

- [1] A. Baumberg and D. Hogg. An efficient method for tracking using active shape models. In *Proc. IEEE Workshop on Motion of Non-rigid Objects*, pages 194–199, 1994.
- [2] Michael Black and Allan Jepson. Recognizing temporal trajectories using the condensation algorithm. In *Proc. International Conference on Automatic Face and Gesture Recognition*, pages 16–21, 1998.
- [3] A. Blake, R. Curwen, and A. Zisserman. A framework for spatiotemporal control in the tracking of visual contours. *International Journal of Computer Vision*, 11:127–145, 1993.
- [4] J. Davis and A. Bobick. The representation and recognition of action using temporal templates. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 928–934, 1997.
- [5] W. Gilks, S. Richardson, and D. Spiegelhalter. *Markov Chain Monte Carlo in Practice*. Chapman and Hall, 1996.
- [6] Michael Isard and Andrew Blake. Condensation – conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29:5–28, 1998.
- [7] Michael Isard and Andrew Blake. Icondensation: Unifying low-level tracking in a stochastic framework. In *Proc. 5th European Conference on Computer Vision*, volume 1, pages 893–908, 1998.
- [8] N. Johnson, A. Galata, and D. Hogg. The acquisition and use of interaction behaviour models. In *Proc. CVPR*, pages 866–871, 1998.
- [9] N. Johnson and D. Hogg. Learning the distribution of object trajectories for event recognition. *Image and Vision Computing*, 14:609–615, 1996.
- [10] D. Magee and R. Boyle. Building shape models from image sequences using piecewise linear approximation. In *Proc. BMVC*, pages 398–408, 1998.
- [11] D. Magee and R. Boyle. Building class sensitive models for tracking application. In *Proc. BMVC*, pages 594–603, 1999.
- [12] A. Pentland and B. Horowitz. Recovery of nonrigid motion and structure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:730–742, 1991.
- [13] L. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77:257–286, 1989.
- [14] N. Sumpter and A. Bulpitt. Learning spatio-temporal patterns for predicting object behaviour. In *Proc. BMVC*, pages 649–658, 1998.
- [15] D. Terzopoulos and R. Szeliski. Tracking with Kalman snakes. *Active Vision*, pages 3–20, 1992.
- [16] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham. Training models of shape from sets of examples. In *Proc. BMVC*, pages 9–18, 1992.
- [17] Michael Walter, Alexandra Psarrou, and Shaogang Gong. Learning prior and observation augmented density models for behaviour recognition. In *Proc. BMVC*, pages 23–32, 1999.
- [18] C.R. Wren and A.P. Pentland. Understanding purposeful human motion. In *Proc. IEEE International Workshop on Modelling People*, pages 19–25, 1999.