A System for the Behaviour Analysis of Laboratory Animal Based on Qualitative Modelling

Maja Matetic
Faculty of Philosophy Rijeka
University of Rijeka
Omladinska 14, 51000 Rijeka, Croatia
maja@pefri.hr

Slobodan Ribaric
Faculty of Electrical Engineering and Computing
University of Zagreb
Unska 3, 10000 Zagreb, Croatia
slobodan.ribaric@fer.hr

Abstract - The paper presents a system LABAQM for the analysis of laboratory animal behaviour based on qualitative modelling. We are dealing with the cognitive phase of the laboratory animal behaviour analysis as a part of the pharmacological experiments. The system is based on the quantitative data from the tracking application and incomplete domain background knowledge. The LABAQM system operates in two main phases: behaviour learning and behaviour analysis. The learning and behaviour analysis phase are based on symbol sequences, obtained by the transformation of the quantitative data. Behaviour learning phase includes supervised learning procedure, unsupervised learning procedure and their combination. The combination of supervised and unsupervised learning procedures produces more robust models of characteristic behaviours which are used in the behaviour analysis phase. The Hidden Markov Model (HMM) has been used as a model of object behaviour, based on qualitative spatial and temporal representation and conceptual grouping method.

I. INTRODUCTION

Tracking of laboratory animal and behaviour interpretation based on the frame sequence analysis has traditionally been quantitative and typically generates large amounts of temporally evolving data. In our work we are dealing with higher-level approach to represent and reason about data such as conceptual clustering [1], [2] and qualitative modelling [3], [4].

The paper deals with additional research in building of the dynamic vision system described in [5], i.e. deals with the problem of the off-line behaviour analysis and recognition of laboratory animal during pharmacological experiments. The quantitative data are obtained by developed tracking system [6]. By using of background knowledge, spatio-temporal model, conceptual clustering and qualitative modelling the animal behaviour analysis and recognition are performed.

During the last two decades, the number of papers within the field of mobile object behaviour capture using computer vision has grown significantly. Aggarwal and Cai [7] describe human motion capture problem as: action recognition, recognition of the individual body parts and body configuration estimation. The surveys of the latest developments in the field are given by Gavrila [8] and Moeslund and Granum [9]. The approaches dealing with recognition are described.

Jonker and Treuer in [10] present their work showing that the study of animal behaviour can benefit from software tools for agent modelling at a conceptual level that support simulation. Generic task models, created by an expert are used to structure the modelling process.

Gavrila in [8] gives an overview of two-dimensional approach that do not consider explicit object shape models. This approach has been especially popular for applications of hand pose estimation in sign language recognition and gesture-based dialogue management. There are several approaches considering the motion trajectories of the hand centroids [11], [12].

Moeslund and Granum [9] predict that field of object motion capture will find inspiration in methods from speech recognition. According to [9] the essential problem is the lack of a general underlying modelling language, i.e. how to map the images into symbols. The work of Bregler [13] introduces such an idea of representing motion data by "movemes" (similar to phonemes in speech recognition). This type of high level symbolic representation is also used in the work by Wren et al. [24]. In their approach they automatically build a "behaviour alphabet" and model, each behaviour using Hidden Markov Models (HMMs).

An overview of the latest research in the field of qualitative spatial representation and reasoning is given by Cohn et al. [4]. Qualitative Spatial Reasoning (QSR) has been used in computer vision for visual object recognition at a higher level which includes the interpretation and integration of visual information. In the work of Fernyhogh et al. [15] QSR technique has been used to interpret the results of low-level computations as higher level descriptions of the scene or video input. Their approach uses qualitative predicates to ensure that scenes which are semantically close have identical or at least very similar descriptions. An example of work using qualitative spatial simulation on the base of conceptual neighbourhood diagrams is the work of Rajagopalan [16].

Our work is an attempt in building a qualitative model for behaviour of laboratory animal. We also use HMM [17], [18] as a model of behaviour.

The paper is organised as follows:
Problem description is given in section II. In section III we describe the LABAQM system for the laboratory animal behaviour analysis problem. We present experimental results and conclude with a discussion of system advantages in section IV.

II. THE PROBLEM OF THE OBJECT BEHAVIOUR ANALYSIS

We present the implementation of the laboratory animal behaviour recognition subsystem which is the part of the dynamic vision system. It is based on methods of frame sequence analysis [6] and it deals with the incomplete background domain knowledge.

We are dealing with the cognitive phase of the laboratory animal behaviour analysis as a part of the pharmacological experiments. On the base of the detected object motion and computed trajectory features we have to analyse and recognise animal behaviours. Each of the object trajectories is presented by a sequence of feature vectors. A feature vector describes the position and orientation of the object in a object trajectory point. Feature vectors are obtained from the dynamic vision system [6] (Fig. 1). Object motion during n consecutive frames is presented by an n-tuple of feature vectors:

$$(((x_1,y_1),\Theta_1),((x_2,y_2),\Theta_2),...,((x_{n-1},y_{n-1}),\Theta_{n-1}),((x_n,y_n),\Theta_n)) \quad (1)$$

The behaviour analysis procedures are based on the following assumptions:

1) The subject remains inside the scene
2) None camera motion
3) Only one object in the scene at the time
4) Movements parallel to the camera-plane
5) No occlusion
6) Subject moves on a flat ground plane

III. THE SYSTEM FOR THE LABORATORY ANIMAL BEHAVIOUR ANALYSIS BASED ON QUALITATIVE MODELLING (LABAQM)

In this section we describe the structure of the LABAQM system based on the quantitative data obtained by the dynamic vision system [6] and the incomplete expert knowledge. The LABAQM system for laboratory animal behaviour analysis operates in two main phases: behaviour learning and behaviour analysis. The learning and behaviour analysis phase are based on symbol sequences, obtained by the transformation of the quantitative data.

A. Feature vector transformation

Object motion description obtained by the dynamic vision system represents the input data to qualitative conversion module (Figure 3.1). The background knowledge base consists of chunks of the incomplete expert knowledge about behaviour attributes described by the problem domain expert. To perform the conceptual clustering procedure of the animal behaviour a spatio-temporal model is proposed. The spatio-temporal model depends on space ontology which is a base for the qualitative modelling. Space ontology includes spatial entities and relationships among them. We may consider topology, their size, the distance between them, their orientation or/and their shape [4]. We have decided to use rectangle elementary regions which are adapted to the object dimensions.

Each n-tuple of feature vectors is divided into k subsequences (k<<n) of equal length. k depends on the expert choice and is expressed implicitely by T, where T is the time duration of a subsequence. Each of these subsequences is converted to a sequence of symbols which describes spatio-temporal ontology. T is n/(k*fps), where fps is frames per second.

1) The spatio-temporal model: The information provided from tracking applications is, by nature, quantitative and described by n-tuple of feature vectors (as in (1)) for the object in the scene. We use the approximate regions, which are conceptually relevant to the physical structure of the domain. For our problem the model of space and time is the base for the feature vector transformation. Qualitative region and qualitative orientation conceptual neighbourhood are given in Figures 2 and 3 respectively. The prime rectangle is chosen as the proper spatial extension of object projection in 2D (Fig. 4), so the object cannot change behaviour qualitative state inside this rectangle (d_1 and d_2 determinate the dimensions of the prime rectangle).

In Figure 5 (a) and (b) we can see two examples of feature vector transformation to a sequence of symbols.

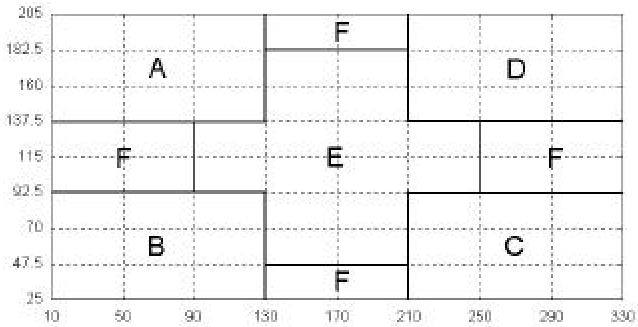


Fig. 2. Qualitative regions

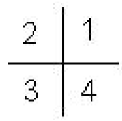


Fig. 3. Set of relations {1,2,3,4} representing second granularity level for orientation

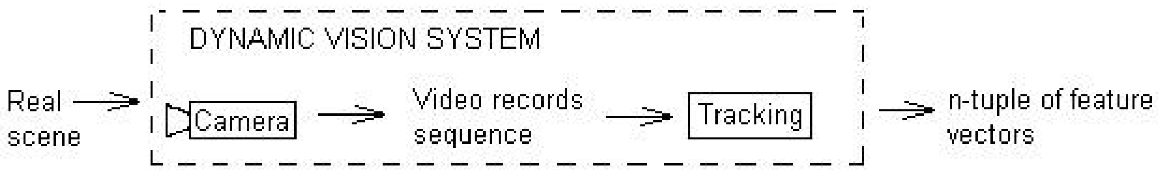


Fig. 1. The dynamic vision system

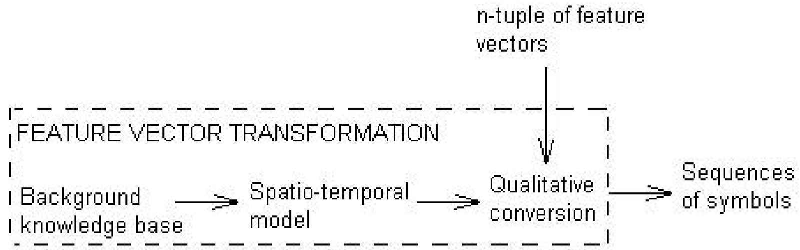


Fig. 2. Feature vector transformation

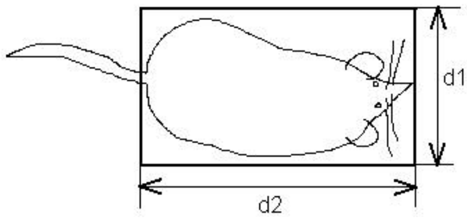
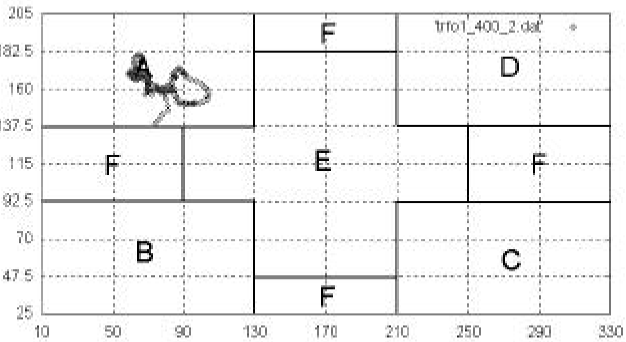
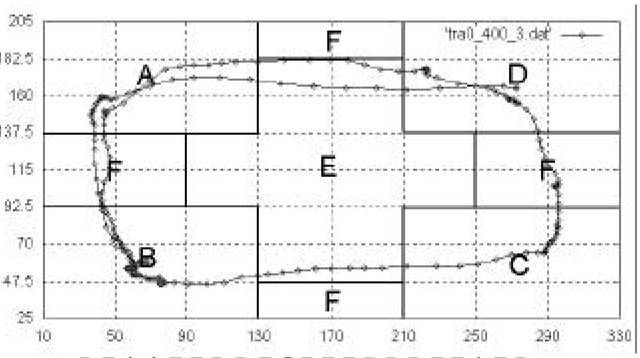


Fig. 4. Spatial extension of the laboratory animal



1. A A
(a)



1. B F A A E E D D F C E E E B B B F F A E D
(b)

Fig. 5. (a) and (b): Examples of feature vector transformation to a sequence of symbols

B. Behaviour learning phase

Behaviour learning phase comprises three subphases: Supervised learning, unsupervised learning and their fusion. Supervised learning includes feature vector transformation subphase and characteristic behaviour modelling subphase (Fig. 6). Unsupervised learning includes feature vector transformation subphase and conceptual clustering subphase (Figures 3.6). Because of the insuficient background knowledge given by the expert and the assumptions made by the motion capture system we propose a combination of supervised and unsupervised learning procedures. The combination of supervised and unsupervised learning produces more robust models of characteristic behaviours which are used in the behaviour analysis phase.

1) Supervised learning: An expert has to choose inserts of video sequences representing characteristic behaviours. These video sequences are inputs to the dynamic vision system [6]. The dynamic vision system generates m-tuples of feature vectors of characteristic behaviours, where m<n. An m-tuple of feature vectors are then transformed in feature vector transformation subphase resulting with sequences of symbols.

According to qualitative region and orientation conceptual neighbourhood in a qualitative state, the qualitative change is restricted to neighbourhood regions (Fig. 2 and 3). Instantaneous transitions depend on preceding transition. This transition interdependence can be built in the modelling procedure. A characteristic behaviour model is created as the hidden Markov model (HMM) [17], [18] for the obtained sequences of symb ols (Fig. 6).

The Hidden Markov Model is a finite set of states, each of which is associated with a probability distribution. Transitions among the states are governed with a set of probabilities called transition probabilities. In a particular state an outcome or observation can be generated, according to the associated probability distribution. It is only the outcome, not the state visible to an external observer and therefore states are "hidden" to the outside; hence the name Hidden Markov Model.

In order to define an HMM completely, following elements are needed: the number of states of the model *N*, the number of observation symbols in the alphabet *M*, set of state transition probabilities $A=\{a_{ij}\}$, where $a_{ij}=p\{q_j \text{ at } t+1|q_i \text{ at } t\}$, observation symbol probability distribution in state j $B=\{b_j(k)$, $b_j(k)=p(v_k \text{ at } t|q_j \text{ at } t)\}$ and initial state distribution $\pi=\{\pi_i\}$, $\pi_i=p(q_i \text{ at } t=1)$. It is assumed that the nnumberext state is dependent only on the current state. This is called the Markov assumption and the resulting model becomes actually a first order HMM.

The characteristic animal behaviours chosen by the expert are modelled as HMMs. We have used HMMs with number of states N ranging from 3 to 10, depending on characteristic behaviours. In our example the expert has chosen two types of characteristic behaviours:

1 - Cycling motion and
2 - Slow motion

The time duration of video sequences used in HMMs training is 40 to 50 minutes. Trained HMMs are used in marking of symbol sequences obtained in unsupervised learning phase.

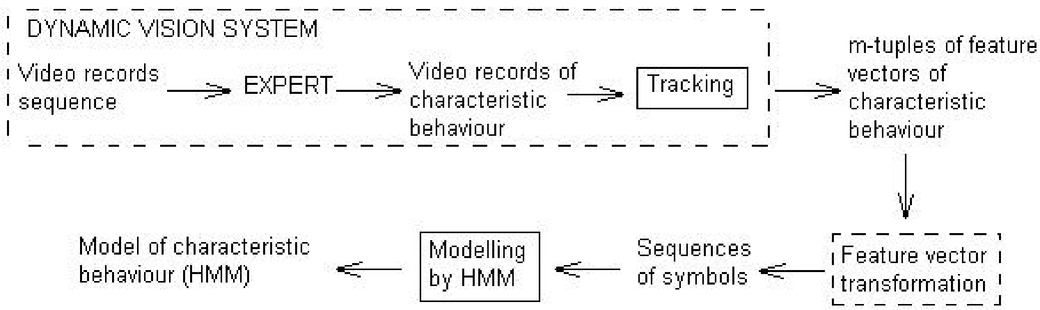


Fig. 6. Behaviour supervised learning

2) Unsupervised learning: The number of symbol sequences, used in HMM training in the supervised learning phase, is too small to be used as a robust model comprising a characteristic behaviour. Larger sets of symbol sequences representing the characteristic behaviours can be obtained by the unsupervised learning method, such as conceptual clustering [1], [2]. At this subphase (Fig. 7) the hierarchical clustering method [19] is used in order to find clusters of input sequences of symbols which are obtained from the feature vector transformation subphase. The similarity measure used for the clustering procedure is based on the Levenshtein distance [20].

Table I gives an overview of the obtained results of clustering procedure for treated (tr_0-tr_3 observation type) and nontreated (ntr_0-ntr_3 observation type) laboratory animals.

In [5] we show that clusters resulting from the unsupervised learning cannot be assumed as learned characteristic behaviours. It is necessary to use the domain expert knowledge in order to recognise groups representing characteristic types of behaviour.

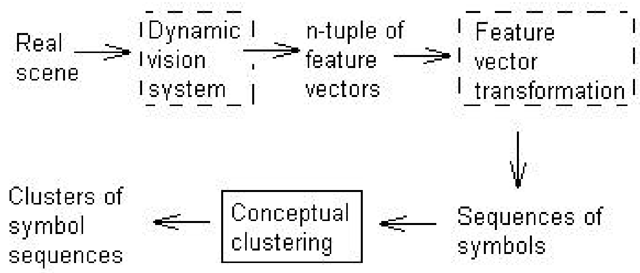


Fig. 7. Behaviour unsupervised learning

TABLE I
NUMBER OF CLUSTERS AS THE RESULT OF CLUSTERING FOR TREATED AND NONTREATED ANIMALS

Laboratory animal class (tr-treated, ntr-nontreated)	Number of significant clusters / total number of clusters
tr_0	2 / 22
tr_1	2 / 23
tr_2	2 / 2
ntr_0	1 / 10
ntr_1	1 / 9
ntr_2	2 / 17

3) The combination of supervised and unsupervised learning: The task of the recognition of characteristic animal behaviours in the scene is to find the connection between the characteristic animal behaviours chosen by an expert, as a part of the background knowledge, and clusters of sequences of symbols obtained by conceptual clustering. Unsupervised learning implemented as conceptual clustering produces clusters containing symbol sequences. It is necessary to use the expert knowledge in order to recognise clusters representing characteristic types of behaviour. So sequences of symbols from clusters obtained by conceptual clustering are analised by the HMMs of characteristic behaviours from supervised learning phase. In this way characteristic behaviours that appear in the scene are recognised and annotated (Table II).

Symbol sequences representing characteristic behaviours chosen by an expert are unified with symbol sequences from clusters obtained in unsupervised phase describing similar behaviours. Characteristic behaviour models are obtained by HMM modelling of these unified symbol sequences (Fig. 8.). These behaviour models are used in the behaviour analysis.

C. Behaviour analysis

HMMs obtained by the combination of supervised and unsupervised learning are used in the process of discovering of characteristic behaviours in the scene (Fig. 9). HMMs are trained for unified symbol sequences representing characteristic behaviours and those obtained by unsupervised learning procedure. An unknown video sequence is marked as recognised behaviour of certain type if the HMM of that behaviour type gives the greater probability P_{max}.

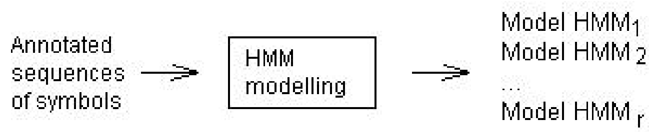


Fig. 8. Combination of supervised and unsupervised learning

TABLE II
ANNOTATED SYMBOL SEQUENCES FROM CLUSTERING PROCEDURE

Observation, Cluster	tr_0 1	tr_0 2	tr_1 1	tr_1 3	ntr_0 1	ntr_1 1
Behaviour type (1-Cycling, 2-Slow motion)	1	1	1	1	2	2

Table III gives the results for the behaviour analysis of video sequence from observations of type tr_3 and ntr_3 (time duration of 55 minutes). Depending on the percentage of marked sequences, these behaviours are recognised as cycling and slow motion respectively. The recognition of unknown behaviour in the case of treated and untreated animal gives good results which have been confirmed by the expert.

IV. CONCLUSION

This work is focusing on the discovery of the qualitative model for the observed behaviour in the scene. Discovered qualitative models is the basis in the analysis of object behaviour.

The presented system LABAQM has following advantages:

- The spatio-temporal model even with incomplete background knowledge is the base for the process of "behaviour alphabet" and automatic qualitative model building.
- It uses HMM as the qualitative model for the object qualitative behaviour in the process of recognition and analysis of object behaviour.

We can compare our system with similar system [10] dealing with the behaviour model building, as the base for simulation of animal behaviour. It uses compositional multi-agent system design method DESIRE and build a simulation model for animal behaviour at conceptual level in an agent-based manner. DESIRE uses generic task models to structure the modelling process. The design of generic models is based on expert explanations so there is much of work preceding the building of a simulation model. Our system LABAQM uses incomplete expert knowledge in the process of automatic qualitative behaviour model building. An HMM can be the base for the simulation of animal behaviour.

TABLE III
BEHAVIOUR ANALYSIS EXPERIMENTAL RESULTS

Observation	tr_3	ntr_3
Behaviour type	Cycling motion (93 %)	Slow motion (85 %)

In our future work we intend to make further system extensions, with the wide scale of possible specific requirements.

V. REFERENCES

[1] M. Kubat, I. Bratko, and R.S. Michalski, *A Review of Machine Learning Methods*, in Michalski, R.S., Bratko,I. and Kubat, M.(Eds), Machine Learning and Data Mining: Methods and Applications, London, John Wiley & Sons, p.p. 3-69,1998

[2] R.S. Michalski, and K.A. Kaufman, *Data Mining and Knowledge Discovery: A Review of Issues and a Multistrategy Approach*, in Michalski, R.S.,Bratko,I. and Kubat,M.(Eds), Machine Learning and Data Mining: Methods and Applications, London, John Wiley & Sons, p.p. 70-112,1998

[3] A.G. Cohn, B. Bennett, J. Gooday and N.M. Gotts, *Representing and Reasoning with Qualitative Spatial Relations about Regions*, in Spatial and Temporal Reasoning, Oliviero Stock (Ed.), Kluwer Publishing Company, pp. 97-134, 1997.

[4] A.G. Cohn, and S.M. Hazarika, "Qualitative Spatial Representation and Reasoning: An Overview", *Fundamenta Informaticae* 43(2001), IOS Press, pp.2-32

[5] M. Matetic and S. Ribaric, "Qualitative Modelling Of Object Behaviour In The Dynamic Vision System Using Hidden Markov Models", *To appear in the Proceedings of the 11th Mediterranean Electrotechnical Conference MELECON 2002*, May 7-9 2002, Cairo, Egypt

[6] Z. Kalafatic, S. Ribaric and V. Stanislavljevic, "A system for Tracking Laboratory Animals Based on Optical Flow and Active Contours", *11th International Conference on Image Analysis and Processing (ICIAP 2001)*, Palermo, Italy, 2001, pp. 334-339.

[7] J.K. Aggarwal and Q. Cai, "Human Motion Analysis: A Review", *Computer Vision and Image Understanding*, 73(3), 1999.

[8] D.M. Gavrila, "The Visual Analysis of Human Movement: A Survey", *Computer Vision and Image Understanding*, vol.73(1), pp.82-98,1999

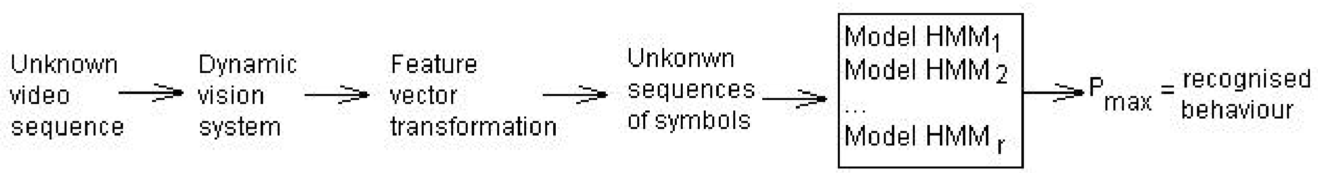


Fig. 9. Behaviour analysis phase

[9] T. Moeslund and A. Granum, "A Survey of Computer Vision Based Human Motion Capture", *Computer Vision and Image Understanding*, 81(3), March 2001.

[10] C.M. Jonker and J. Treuer, "Agent-Based Simulation of Animal Behaviour", *Applied Intelligence* 15, pp. 83-115, 2001

[11] T. Starner, J. Weaver and A. Pentland, "Real-Time American Sign Language Recognition Using Desk and Wearable Computer Based Video", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.20., No. 12, December 1998, pp. 1371-1375

[12] C. Vogler and D. Metaxas, "Parallel Hidden Markov Models for American Sign Language Recognition", *Proceedings of the ICCV(2)* 1999: 116-122

[13] C. Bregler, "Learning and Recognizing Human Dynamics in Video Sequences", *In Conference on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, 1997.

[14] C.R. Wren, A. Azarbayejani, T.Darell and A. Pentland, "Pfinder: Real-time tracking of the human body", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):780-785,1997.

[15] J. Fernyhough, A.G. Cohn and D. Hogg, "Constructing qualitative event models automatically from video input", *Image and Vision Computing*, 18, 2000, pp. 81-103.

[16] R. Rajagopalan, "A model for integrated qualitative spatial and dynamic reasoning about physical systems", *Proceedings of American Conference on AI (AAAI-94)*, 1994, pp.1411-1417.

[17] L.R. Rabiner, B.H. Juang, "An Introduction to Hidden Markov Models", *IEEE ASSP MAGAZINE*, 1986, pp. 4-16

[18] L.R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", *Proceedings of the IEEE*, Vol.77, No.2, February 1989.

[19] G.N. Lance and W.T. Williams, "A General Theory of Classification Sorting Strategy", *Computer Journal*, No. 9, 19-73, 1967.

[20] R.A. Wagner and M.J. Fisher, "The String to String Correction Problem", *Journal of the Association for Computing Machinery*, 21(1), 168-173, 1974.