Qualitative Modeling of Object Behaviour in the Dynamic Vision System

Maja Matetic
Department of informatics
Philosophical Faculty Rijeka
Omladinska 14, 51000 RIJEKA, Croatia
Telefon:051-345 046 Fax:051-345 207 E-mail:maja@mapef.pefri.hr

Abstract - The subject of this paper is cognitive analysis of the dynamic scene: on the base of the identified motions of objects and calculated trajectory attributes we need to analyse, evaluate and reconstruct behaviours implementing artificial intelligence methods. Dynamic scene analysis has traditionally been quantitative and typically generates large amounts of temporally evolving data. Recently, increasing interest has been shown in higher-level approaches to representing and reasoning with such data using conceptual and qualitative approaches. One motivation for the present work was the desire to apply qualitative spatio-temporal reasoning techniques to real-world dynamic scene analysis. The system presented in the paper uses the qualitative modelling method, based on the model of space and time and incomplete domain background knowledge, in the process of discovering qualitative behaviour patterns. The resulting qualitative model of the dynamic scene is the base for the comparative analysis, data interpretation and system simulation.

I. INTRODUCTION

Dynamic scene analysis has traditionally been quantitative and typically generates large amounts of temporally evolving data. Recently, increasing interest has been shown in higher-level approaches to representing and reasoning with such data using conceptual and qualitative approaches [4], [7]. These methods are the part of the system I am proposing in this paper.

This paper deals with the problem of the interpretation of object behaviour in the scene. Problem description is given in section II. In section III I describe the system which can be implemented as the solution for the object behaviour analysis problem, based on the quantitative data from the existing tracking application. The importance of the qualitative reasoning in making conclusions and predictions on the system behaviour, even without complete data, makes it suitable for many real world problems. Unsupervised data classification includes different methods for discovering of natural groups in multidimensional data based on the measured or observed samples similarity. Section IV introduces the model of space and time and the first phase of data grouping procedure. The second phase of the data grouping procedure is described in section V. Characteristic behaviours from background knowledge are modelled and used in the recognition process of the interesting behaviours observed in the scene (sections VI and VII). Qualitative behaviour tree obtained by the qualitative simulation and marked in the behaviour recognition phase is the base for the generation of behaviour explanations and predictions.

II. THE PROBLEM OF THE OBJECTS BEHAVIOUR ANALYSIS AND INTERPRETATION

The problem this work is focusing on is the development of the part of the dynamic vision system which has the theoretical basis in implemented methods of the frame sequence analysis and the use of the knowledge base in the subjects motion and behaviour interpretation [9], [10].

More precise I am dealing with the cognitive phase of the object behaviour analysis: on the base of the detected object motion and computed trajectory attributes, implementing artificial intelligence methods, we need to analyse, evaluate and predict behaviours. Each of the object trajectories is the sequence of attributes vectors. Attributes vectors describe position and orientation information in each of the trajectory points during objects motion. Existing object tracking application generates, for the object present in the scene for n consecutive frames, the sequence T_i of n 2D picture coordinates in equal time intervals:

$$T_i = \{(x_1,y_1),(x_2,y_2),(x_3,y_3),\ldots,(x_{n-2},y_{n-2}),(x_{n-1},y_{n-1}),(x_n,y_n)\} \qquad (1)$$

The trajectory can also be described in term of the flow vector represented with the object position and its velocity:

$$f=(x,y,v) \qquad (2)$$

III. THE SYSTEM FOR THE OBJECTS BEHAVIOUR ANALYSIS BASED ON QUALITATIVE MODELLING

In this section I propose a solution for solving the problem of the objects behaviour analysis, based on the quantitative data obtained by the existing tracking process [9], [10] and the background knowledge. The background knowledge consists of the behaviour attributes and the set of characteristic behaviours described by problem domain expert.

The system for the behaviour analysis consists of several main modules (figure 1):

- The existing tracking process [10]:

Trajectory descriptions obtained by the tracking process are the input data for the quantitative-to-qualitative conversion and conceptual grouping module. More detailed description of the input data is given in section II.

- Background knowledge:

In order to guide the process of the unsupervised conceptual grouping, the first step in creating of the space-time model is the choice of the proper space ontology as the base for the qualitative modelling. The researchers in the field of qualitative modelling and reasoning in 2D space have accepted the region as the prime entity of space, and reasoning is based mostly on transitivity tables which guide a simulation. After the space ontology is chosen, domain expert gives attribute values to the qualitative regions in the scene of the interest (section IV). Attributes are given also to the objects in the scene describing their physical properties (for example: treated or nontreated laboratory animals). The expert gives also the set of characteristic behaviours, which are used in the process of interesting behaviour recognition (section VI).

- Quantitative-to-qualitative behaviour conversion:

Quantitative-to-qualitative conversion of the tracking data is based on the spatio-temporal model. Complex quantitative behaviour is divided into several simple quantitative behaviours of equal time duration Δt and these are the input quantitative samples. Each of these samples is converted to the qualitative behaviour on the base of spatial and temporal ontology and depending on expert choices of the attributes. The new qualitative state is added to the qualitative behaviour when a qualitative change of some attribute is detected (section IV).

- The unsupervised behaviour conceptual grouping:

This module determines the similarity of qualitative behaviours and includes also the hierarchical grouping method [8] (section V).

- Characteristic behaviours modelling:

The task of the recognition of interesting behaviours is to find the connection between the characteristic qualitative behaviours given by an expert and qualitative behaviour groups obtained by conceptual grouping procedure. A characteristic behaviour model can be created as a hidden Markov model (HMM) through the training procedure, which can be used later in the process of interesting qualitative behaviour recognition (section VI).

- Marking the behaviour groups:

Qualitative behaviour group prototypes, as the result of conceptual grouping procedure, can be sent to the input of the HMM which is the stochastic finite automata [8]. In this way interesting behaviours that appear in the scene are detected (section VI).

- Qualitative spatial reasoning:

On the base of the qualitative spatial model and the initial qualitative behaviour state, which is given by an expert through his interface, the tree of the qualitative behaviours is formed. The tree branches are then marked with the recognized characteristic behaviours. The marked behaviour tree is the base for the behaviour explanation generation and predicting of the future qualitative behaviour states (section VII).

- Expert interface:

The possibility of event reconstruction using the simulation and predicting of the future system states is the base for development of the tutoring systems for teaching experts (for example in tracking of laboratory animals in pharmacology or in tracking of billiard-balls).

IV. THE SPATIO-TEMPORAL MODEL AND QUANTITATIVE-TO-QUALITATIVE BEHAVIOUR CONVERSION

The principal goal of Qualitative Reasoning (QR) is to represent not only everyday commonsense knowledge about physical world, but also the underlying abstractions used by experts when they create quantitative models [1]. Endowed with such knowledge, and appropriate reasoning methods, a computer could make predictions, diagnoses and explain the behaviour of physical system in a qualitative manner, even when a precise quantitative description is not available. The key to a qualitative representation is not simply that it is symbolic, and utilizes discrete quantity spaces, but that distinctions made in these discretisations are relevant to the behaviour being modelled - i.e. distinctions are only introduced if they are necessary to model some particular aspect of the domain with respect to the task in hand.

Traditionally, in mathematical theories of space, points are considered as primary primitive spatial entities, and extended spatial entities such as regions are defined as sets of points. However, within the qualitative spatial reasoning community, there has been a strong tendency to take regions of space as the primitive spatial entity. Topology is perhaps the most fundamental aspect of space and it must form a fundamental aspect of qualitative spatial reasoning since topology certainly can only make qualitative distinctions.

Orientation is a naturally qualitative property. If we want to specify the orientation of a primary object with respect to a reference object, then we need some kind of frame of reference. An extrinsic frame of reference imposes an external orientation. A deictic frame of reference is with respect to the "speaker" or some internal observer. Finally, an intrinsic frame of reference exploits some inherent property of the reference object.

The information provided from existing tracking applications is, by nature, quantitative with the position and spatial extent of objects usually provided in screen coordinates. However, using the approximate zone or region rather than the exact location will collapse broadly similar behaviours into equivalence classes to provide a generic model. Of course a scene cannot be arbitrarily segmented into regions - rather, the regions should be conceptually relevant to the physical structure of the domain rather than arbitrary. An existing approach to the problem of automatic qualitative modelling of events in the scene is described in [2].

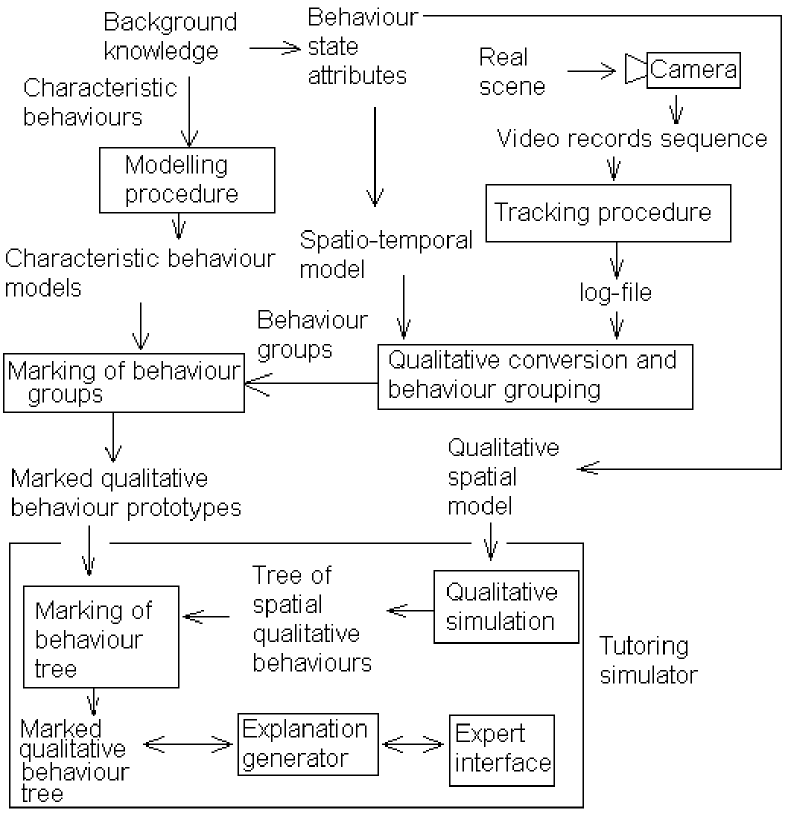


Figure 1: Qualitative modelling procedure and tutoring simulator as the part of the dynamic vision system

For our problem the model of space and time is the base for the quantitative-to-qualitative conversion. Qualitative behaviour is the sequence of qualitative states each of them represented with n attribute values. The attributes acquire their values from the given domains. The conceptuality of the grouping process is achieved by choosing qualitative attribute, which reflects the relationship between elementary attributes. It is the case of the qualitative region that reflects the relationship between two attributes for x and y coordinates.

In first step we deal with the topology view to the problem. It is to describe first attribute, which is the qualitative region. To make easier the detection of the qualitative change in space, a rectangle mesh can be used (although the region transitivity tables are not limited to that choice of topology and depends only on qualitative region change detection procedure). The prime rectangle is chosen as the proper spatial extension of object projection in 2D, so the object cannot change behaviour qualitative state inside this rectangle (for cause of orientation or velocity change). The expert can choose values for qualitative region attribute by marking those prime rectangles by symbols from the finite alphabet (for example look at the figure 2 where K1-K4 represents corner, S center and R border qualitative region). Again in order to simplify possible transitions between qualitative regions, we can assume that only neighbourhood prime rectangles can be marked with the same symbol. The prime rectangles are then unified resulting in new topology frame where qualitative regions are of different shapes and sizes. The transitivity table describes possible transitions between these qualitative regions. Table 1 shows possible transitions for qualitative region values in figure 2. The finite alphabet and marked qualitative regions form a part of the background knowledge.

The second step is to define the second attribute, which is the shape of the object trajectory. It is considered inside instantaneous qualitative region and is defined through the change of the orientation as it is shown in figure 3(a,b) [8]. An extrinsic frame of reference is used. Background knowledge contains also values for trajectory qualitative shape attribute, for example: (p - passing through, c - cycling), which has its own description built from elementary symbols.

Time attribute of the qualitative state is qualitative state time duration. We can choose ordinal domain of values, for example: (s - short, n - normal, l - long). These values are supplied also from background knowledge, and must be defined as quantitative value intervals by the expert. The shape and time attribute qualitative values specify qualitative states in duration and trajectory shape. This information is added in form of indices to the symbols, which represent qualitative region of the behaviour state. In this way the former alphabet representing different qualitative regions is extended to the new one V where every symbol representing different qualitative region has indices representing time and shape attribute value.

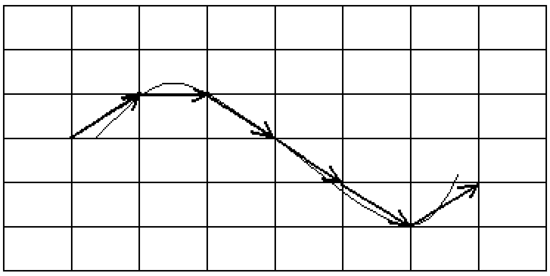
Quantitative-to-qualitative conversion is based on the spatio-temporal model and the log-file. It produces the set of qualitative behaviours. Complex quantitative behaviour is split in n simple quantitative behaviours of equal time duration Δt. Each of these quantitative samples is

K4	K4	R	R	R	R	K3	K3
K4	K4	R	R	R	R	K3	K3
R	R	R	S	S	R	R	R
R	R	R	S	S	R	R	R
R	R	R	S	S	R	R	R
R	R	R	S	S	R	R	R
K1	K1	R	R	R	R	K2	K2
K1	K1	R	R	R	R	K2	K2

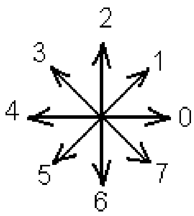
Figure 2: Elementary and complex qualitative 2D regions

TABLE I
TRANSITIVITY TABLE FOR QUALITATIVE REGION ATTRIBUTE

Instantaneous qualitative region	Possible transitions
K1	R
K2	R
K3	R
K4	R
R	K1, K2, K3, K4, S
S	R



(a)



(b)

Figure 3 (a,b): Trajectory description by sequence of symbols representing elementary samples

converted to the qualitative behaviour on the base of spatio-temporal model and behaviour state attribute values.

V. QUALITATIVE BEHAVIOURS CONCEPTUAL GROUPING

Quantitative-to-qualitative conversion is a kind of grouping method, which converts a sequence of attribute vectors of time duration Δt to a string of symbols from the finite alphabet. The alphabet is determined by the choice of the spatial and temporal ontology and the background knowledge. Next phase of grouping includes computing of similarity between pairs of strings (qualitative behaviours) using Levenshtein distance [8]. These strings of symbols are of the different length so to determine similarity the dynamic programming method can be used.

Let V* be the set of all strings of symbols from the finite alphabet V (described in section IV). Levenshtein distance between two strings of symbols x and y from V^* is the minimum number of symbol mappings which translate string x to string y. Possible mappings are:

1. Exchange symbol

$\alpha a\beta \rightarrow \alpha b\beta,\ \forall a,b \in V,\ a \neq b,\ \alpha,\beta \in V^*$

2. Delete symbol

$\alpha a\beta \rightarrow \alpha\beta,\ \forall a \in V,\ \alpha,\beta \in V^*$

3. Insert symbol

$\alpha\beta \rightarrow \alpha a\beta,\ \forall a \in V,\ a \neq b,\ \alpha,\beta \in V^*$

Levenshtein distance between two strings x and y can be written as:

$D_L(x,y)=\min_j\{Z_j+B_j+V_j\}$, $j=1,2,...,J$, where:

Z_j is the number of exchanged symbols

B_j is the number of deleted symbols

V_j is the number of inserted symbols and

J is the number of possible translations of string x to string y.

Dynamic programming algorithms are used for finding shortest paths in graphs and the comparison or alignment of strings (as in biological DNA, RNA and protein sequence analysis, speech recognition and shape comparison). The cost of the transformation of string x to string y is equal to the sum of the costs of all the symbol mappings from the mapping sequence. To minimize the cost of the transformation we use dynamic programming.

In order to use dynamic programming for the computation of the Levenshtein distance between strings:

$x=a_1a_2...a_n$ and

$y=b_1b_2...b_m$,

we introduce symbols for partial strings:

$x(i)=a_1a_2...a_i$, $i=1,2,...,n$

$y(j)=b_1b_2...b_j$, $j=1,2,...,m$ and

the symbol for Levenshtein distance between partial strings x(i) and y(j):

$D(i,j)=D_L(x(i),y(j))$,

which is computed as:

$D(i,j)=\min\{m1,m2,m3\}$,

where

$m1=D(i-1,j-1)+Z(a_i,b_i)$,

$m2=D(i-1,j)+B(a_i)$ and

$m3=D(i,j-1)+V(b_j)$.

Follows the algorithm for computing the Levenshtein distance between strings x and y:

```
D(0,0):=0;
for i:=1 to n do
   D(i,0):=D(i-1,0)+B(a_i);
end-for;
for j:=1 to m do
   D(0,j):=D(0,j-1)+V(b_j);
end-for;
for i:=1 to n do
   for j:=1 to m do
      m1=D(i-1,j-1)+Z(a_i,b_i);
      m2=D(i-1,j)+B(a_i) ;
      m3=D(i,j-1)+V(b_j);
```

    D(i,j):=min{m1,m2,m3}
  end_for
end_for
Levenshtein distance $D_L(x,y)=D(n,m)$.

Qualitative sample grouping is accomplished by the hierarchical grouping method [8]. It is based on the Levenshtein distance.

VI. CHARACTERISTIC BEHAVIOUR MODELLING AND RECOGNITION OF INTERESTING BEHAVIOURS IN THE SCENE

An expert through the bacground knowledge provides characteristic behaviours which are used in the process of recognition of interesting behaviours in the scene. According to qualitative region transitivity table in a qualitative state the change of qualitative region is restricted to neighbourhood regions. Instantaneous transitions depend on preceding transition. This transition interdependence can be built in the process of recognition by modelling the characteristic qualitative behaviour with hidden Markov chain [8].

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 $\Lambda=\{a_{ij}\}$.

$a_{ij}=p\{q_{t+1}=j|q_t=i\}$, $1<=i,j<=N$

where qt denotes the current state.

It is assumed that the next state is dependent only upon the current state. This is called the Markov assumption and the resulting model becomes actually a first order HMM.

The first step in recognition process is to learn the hidden Markov model for each of the characteristic behaviours from background knowledge. These models are then used in recognizing behaviour groups achieved by the hierarchical grouping procedure. Qualitative behaviours recognized as interesting are then marked in the qualitative behaviour tree obtained by the qualitative simulation procedure.

VII. QUALITATIVE SPATIAL SIMULATION

Although much of the work in qualitative spatial reasoning has concentrated on representational aspects, various computational paradigm are being investigated including constraint based reasoning. The most prevalent form of qualitative spatial reasoning is based on transitivity table (composition table). Perhaps the most common form of computation in the traditional qualitative reasoning is qualitative simulation. Using conceptual neighbourhood diagrams is quite easy to build a qualitative spatial simulator. Figure 4 shows an example of conceptual neighbourhood for RCC8 calculus. Such simulator takes the description of an initial state and constructs a tree of future possible states - the branching of the tree results from the ambiguity of the qualitative calculus. Figure 5 shows tree levels of qualitative behaviour tree obtained for initial state with qualitative region value K1 and the transitivity table (table 1) for topology aspect of space.

Qualitative behaviours, which are recognized as interesting in the process of characteristic behaviour recognition, are marked in the qualitative behaviour tree adding information about shape and time attribute values.

One important problem is the need for intelligent tutoring systems and learning environments for expert education and training [3]. Tutoring simulators can explain as well as reproduce the behaviour of what they are modelling, and thus provide a basis for deeper reasoning about behaviour. In my previous work I developed such a tutoring system by combining numerical and qualitative simulations [5],[6].

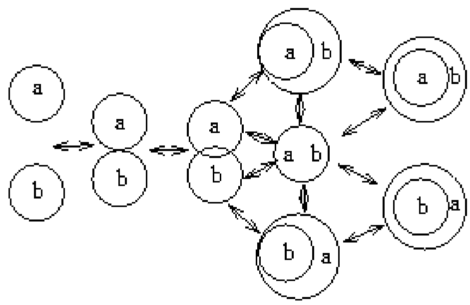


Figure 4: Conceptual neighbourhood in RCC8

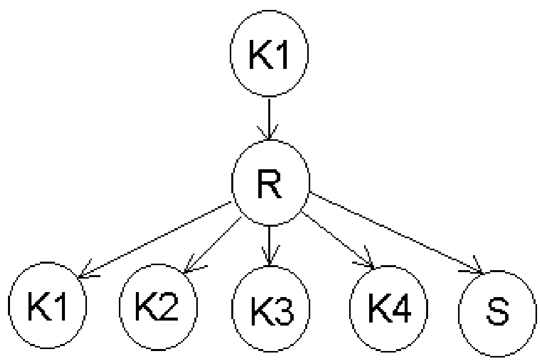


Figure 5: Three levels of the qualitative behaviour tree for given initial state and transitivity table (table 1)

In our present problem marked behaviour tree in combination with quantitative data is the base for the generation of explanation and prediction of future object behaviours. There is a class of questions that expert can ask about the system:

1. Summarise the behaviour
2. What is happening at t=..?

3. What happens next?
4. What else might have happened?

VIII. CONCLUSION

This work is focusing on the discovering of the qualitative model for the observed behaviour in the scene. Discovered qualitative models and qualitative spatial simulation are the basis in reasoning about the object behaviour. Further research will be concerned with the directing and refining the unsupervised grouping procedure, depending on the success of model discovering.

LITERATURE

[1] 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.

[2] J.Fernyhough, A.G.Cohn and D.C.Hogh, Building Qualitative Event Models Automatically from Visual Input, Sixth International Conference on Computer Vision, ICCV98, pp. 350-355, 1998.

[3] K.D. Forbus, Using Qualitative Physics to Create Articulate Educational Software, IEEE Expert, Vol. 12, No.3: May/June 1997, pp.32-41

[4] Kubat,M.,Bratko,I. and Michalski, R.S, 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

[5] M.Matetić, "MARK - A Simulation System for Parameter Estimation in Simulation of Ecological Processes", *The 12th International Conference on Case Method Research and Case Method Application: Environmental Problem Solving - From Cases and Experiments to Concepts, Knowledge, Tools and Motivations*, Maribor, Slovenia, November 12-15, 1995, Rainer Hampp Verlag, München, Mering, 1997.

[6] M.Matetic, Explaining Dynamic Systems by Combining Qualitative and Numerical Simulation, M.Sc. Thesis, Faculty of Electrical Engineering and Computer Science, Ljubljana, 1995.

[7] Michalski, R.S. and Kaufman,K.A., 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

[8] N.Pavešić, Raspoznavanje vzorcev, Uvod v analizo in razumevanje vidnih in slušnih signalov, Založba FE in FRI, Ljubljana 2000.

[9] S.Ribarić, D.Gluhak, Computer Vision System for Tracking Coloured Objects, Proceedings of MIPRO'97 conference, Opatija 1997.

[10] S.Ribarić, C.Dodoczky, Z.Kalafatić, V.Stanisavljević, Real-Time System for Analyzing Animal Motion, Proceedings of MIPRO'97 conference, Opatija 1997.