TEMPORAL EVENT RECOGNITION:

AN APPLICATION TO LEFT VENTRICULAR PERFORMANCE

John K. Tsotsos

Department, of Computer Science.
and
Division of Cardiology, Toronto General Hospital
University of Toronto
Toronto, Ont., Canada

ABSTRACT

The problem of recognition of time-varying events, particularly motion concepts, is discussed and a framework for the abstraction of such concepts is briefly presented. We have designed and implemented a knowledge-based expert system for motion understanding that incorporates several novel features. A frame-based representation, which includes exception handling via similarity links and the organizational primitives IS-A and PART-OF is used to construct a knowledge base of temporal concepts. This knowledge base drives the recognition process that integrates the paradigms of hypothesize-and-tr»st and competition and co-operation among conceptually adjacent hypotheses. The recognition stra tegy also includes a hypothesis rating scheme that is amenable to empirical performance analysis. This framework has been applied to the problem of human left ventricular performance evaluation from image sequences, and thus the knowledge base also contains much of the current understanding of left ventricular dynamics. This paper describes some aspects of this application in detail and presents a complete analytic of a typical image sequence, emphasizing comparisons between human and computer analysis. Additional justification of the adequacy of the framework has been obtained using empirical studies of the hypothesis rating scheme.

INTRODUCTION

The problems of recovering three dimensional structure of moving objects from an image sequence along with the coi responding translational and rotational parameters, and of interpreting "apparent" motions, have received a substantial amount of attention in the recent past [1], [2], and [3]. Uur approach is complementary to these in that we address the questions of 'In what *context* is this motion information useful and in what *form,* is it useful, and how can a computer system *abstract* the motion summary from the large amount of quantitative data?". Many applications for motion analysis are described in [4], and include outdoor scenes, clinical imagery, and satellite and air born imagery. In particular, two and three-dimensional imagery of the dynamic left ventricle (LV) are becoming increasing popular and useful as clinical aidB to diagnosis of disorders of the heart. The left ventricle is perhaps the most important component of our heart since it supplies blood to the body. Yet, its analysis is performed mostly by human observation of its imagery with little or no computer assistance, even though the motions exhibited by the LV are more complex than the human visual system can objectively and consistently analyse.

This "real world" application U rich with temporal concepts and relationships and motivated our research into time-varying event understanding.

Our system, called ALVKN, is a knowledge-based expert system for human left ventricular performance assessment. Input is an image sequence obtained by cineradiography of the human left ventricle However, the methodology is applicable to other forms of dynamic cardiac imagery as well. The goal of the ALVEN system is two-fold: to experiment with and study motion understanding methodologies in the context of a rich temporal domain; and, to consistently and objectively evaluate the performance of the left ventricle and in doing so. compile and refine the relevant cardiology knowledge, with the intent of providing both a clinical and research tool for cardiologists.

We will briefly describe the interesting features of our representation formalism and recognition control structure, and then give details about a particular analysis of a lfft ventricular film.

A FORMALISM FOR REPRESENTING TIME-VARYING EVENTS

In two quite comprehensive collections of current time-varying imagery research. [4], [5], the issue of representation of t.prnpor«l information, particularly for motion, seems quite neglected. Yet. it is clearly an important issue, especially for a motion understanding framework such as ours.

The main concern of our representational formalism is how temporal concepts are to be represented and organized into a knowledge base (KR) The intent is to provide a formalism for representing general temporal concepts* such that problem-specific ones, in our case, left ventricular motions, can be represented in terms of the; general concepts-. Furthermore, the form of the KB must be such that it can be used by an appropriate control structure to guide recognition of the problem-specific concepts stored in it. There are several types of temporal concepts that we are concerned with:

a) event start and end states;
b) rates of change during an event;
c) form of the changes during the event:
d) changes *of* relationships of the subject of the event with respect to other events or objects;
c) temporal constraints between events

In addition to providing a coherent scheme for representing such concepts, the formalism must take several other constraints into account:

a) The data involved often have a "fuzzy" quality. In real world domains, there are ranges of acceptable values to contend with, as opposed to single correct values, as well as overlapping ranges of acceptable values between conflicting concepts. This is particularly important for events whose normal time intervals may be overlapping and which exhibit a temporal precedence relationship (event A must precede event B, but their temporal boundary is fuzzy).

b) The representation must have a strong descriptive basis with qualitative concepts related to quantitative ones, that is, there must be a "bridge" between descriptive and numerical concepts.

c) The formalism must be capable of representing concepts at different levels of abstraction, depending on the level of expertise of the person who enters the knowledge. Both types of descriptions must be handled because there are cases where only coarse descriptions are available that must be integrated with more detailed descriptions.

Finally, an investigation into appropriate reognition-biased knowledge structuring primitives is required for the construction of the knowledge base.

Entities of the Representation

The basic entities of the representation are packages of information called *frames* and are used to model abstract temporal concepts such as CONTRACT or EXPAND, and abstract object concepts such as HEART or LV. Each frame may have an associated description of its internal structure that defines the semantic components of the concept that the frame represents. For temporal concepts, the internal structure of a frame defines the semantic components of the event, such as rate or duration. On the other band, object concept frames have an internal structure that defines their physical parts and the properties possessed by them. The internal structure of a frame is defined by *slots,* each of which has a name and an associated *type.* The type relates the slot to other frames that specify the kind of tokens that can fill (be bound to) it.

Slots may have associated information that specifies a default value, constraints that have to be satisfied by its values, and exceptions that are to be raised when these constraints are not satisfied, thus driving the recognition process. Exceptions are instances of the exception type frame and contain all relevant information required by the recognition mechanism to characterize the matching failure, and to assist in determining which alternate hypotheses may be appropriate.

Our representation is closely linked with the *PSN* formalism [6] and the concepts of frames and organizational primitives are borrowed from there. Two primitive relationships are used to organize frames into a knowledge base with different levels of abstraction The *PART-OF* relationship relates a frame to its parts, which are the types of its slots. Clearly, recognition of an instance of a frame makes heavy use of the PART-OF hierarchy in the recognition of its parts. This relationship is a tool useful for controlling the level of detail in a frame, and for providing a bridge between quantitative and qualitative descriptions. For example, at a coarse level of detail, a concept such as CONTRACT is appropriate while at a much finer level, the motions of the individual components that taken together make up the contraction could be found.

The other primitive relationship used to organize frames is called the *IS-A* relationship. It is intended to capture the notion of generalization or classes and sub-classes. The use of IS-A economizes on frame size (through inheritance of properties), and provides a mechanism for imposing more global constraints on specific concepts.

Our representation scheme is clearly biased towards recognition. Perhaps the most important representational feature for the purposes of recognition is the *similarity link* [7]. These links relate frames that, are, in some sense, similar, and are used to handle situations where one or more matching exceptions have been raised. Frames that have an active similarity link between them are assumed to be mutually exclusive. The information associated with a similarity link is stored inside a generic frame. The link remains dormant, i.e., the source frame and the specified destination frame are not connected in any way, until the recognition process deems the connection necessary.

Each similarity link carries two types of information (in addition to its source frame and destination frame(s)):

a) A *condition* which determines whether the raised exceptions can indeed be handled by the link; and

b) *binding information* that indicates how the values of the slots in the source frame are related to those in the destination frame.

The condition is made up of two components, a *similarity expression* and a *difference expression.* The similarity expression contains the necessary conditions which the two linked frames must share in order for the similarity link to be activated (i.e., the "similarities"). Since only one of the frames (the source frame) is active when this condition is evaluated, the predicates describe prerequisite components of the destination frame that must also be possessed by the source frame. The difference expression is a list of exceptions which may be raised during matching of the source frame. Each of the list's elements has an associated time interval that specifies when each exception should occur if all of them are observed in the image sequence. In effect, such an expression indicates the time course of the matching failures in the source frame context, that indicate that the destination frame's context would be more relevant (i.e., the "time sequence of differences"). Thus, the similarity links point to alternate frames potentially viable when exceptions are raised.

Temporal Concepts

The temporal concepts of interest for our purposes are those describing two-dimensional spatial changes. However, the representation formalism is not restricted to only these and, we believe, Is extendable to the three-dimensional case as well. In order to adequately represent motion concepts, certain semantic components must be included. Miller [6] has done a very good analysis of the semantic components for a large class of English motion verbs. Using Miller's work as a foundation, Badler [9] refined these concepts, gave detailed definitions of the directionals, and outlined a framework within which recognition can be done. The work presented here is strongly based on these research efforts. Miller's semantic components include changes of location, the reflexive-objective distinction, direction of motion, changes of physicai properties of an object, changes of motion, rate of motion and others. Our representation can handle these, but currently does not deal with causality.

Motion frames, once defined, are organized into a KB using the IS-A. PART-OF and Similarity relationships. Motions are divided into two main subclasses - simple and aggregate motions. The simple ones are "pure" motions - strict contraction, strict upwards motion, strict lengthening, etc. Motions which are combinations of these simple motions, are classified under the class of aggregate motions. This specialization uses the IS-A relationship. The aggregation of several simple and/or other aggregate motions into an aggregate sub-class is accomplished via the PART-OF relationship. Motions which cannot coexist for the same object during the same time interval are represented by including similarity links in their definitional frames. The activation or use of any of these three relationships is determined during processing by the control structure. Details of the frame definitions for motion classes as well as the organization of motion classes can be found in [10].

THE RECOGNITION SCHEME

Overview

The paradigms of competition and cooperation among hypotheses and hypothesize-and-test form the basis of our recognition control structure. The key feature of the control structure is that it is driven by the organization of the knowledge base, that is, by the primitive relations between knowledge units. Activation of hypotheses proceeds along the IS-A, PART-OF and Similarity relationships in a constrained manner.

Hypotheses are ranked on the basis of certainty factors. Each hypothesis, when activated, receives an Initial certainty factor equal to that of the hypothesis that activates it. A modified relaxation labelling process (RLP) [11] is then used to update the certainty factors. The relaxation process is based on *conceptual adjacency* that specifies which hypotheses are competitors and which ones are complementary and in what respect. The compatibilities between hypotheses that are necessary for the RLP are derived in a dynamic fashion, depending on which conceptual adjacencies are present between hypotheses that are active during the course of the input image sequence. The best hypotheses (highest ranked) are used to derive the expectations for the next image.

A scheme similar to, but not as sophisticated as [12], is proposed for the vision aspects of processing, that is, the use of expectations to guide the search for objects in the image and determine their correspondencies to the previous image. Note that medical researchers have already proposed successful solutions for locating the features in our particular type of imagery [13]. Changes are described in terms of location changes of points, length changes of axes and perimeters, area changes and shape changes. These four primitive kineses provide an intermediate representation for relating quantitative changes to qualitative ones.

This low level data drives the activation of hypotheses that attempt to describe the exhibited motions. The starting set of hypotheses for each object in the first image contains only one hypothesis - the NO MOTION hypothesis. The kineses for a particular object are matched against the hypothesized motions, and matching failures between expected and actual kineses are recorded. These failures are represented in terms of exception frames which contain any information necessary so that proper selection can be made of alternate hypotheses. This selection is made via the similarity links which are present in the frame of which an active hypothesis may be an instance. The activation of more specialized hypotheses proceeds along the IS-A axes in the KB. Selection of an appropriate IS-A child frame depends on the matching success of a frame which airframes have as one of their specializations - the NO SPECIALIZATION frame. This frame inherits all the properties of its IS-A father and adds new constraints to Its definition, so that if it succeeds in matching, it excludes all other specializations of its IS-A father. It also contains a set. of similarity links relating it via possible exceptions to the other specializations of its IS-A father. In this way, if it fails the matching process, other specializations are activated. It also allows the system to identify when there are no frames in the KB that are appropriate for describing the exhibited motion. Finally, aggregate hypothesis frames are activated on the success of a component vin the corresponding PART-OF link. *In* other words, the IS-A relationship is a mechanism by which the frames along o particular "path" in the network are activated, each one being a specialization (generalization) of another frame on the path, the PART-OF relationship allows a breadth-first search for frames corresponding to aggregations of frames perhaps on several different IS-A paths, and Similarity links provide a mechanism for selectively adding the frames of other IS-A paths into the active hypothesis net.

Let us, for a moment, compare this with the change of attention mechanisms in two other large application systems HEARSAY-I1 [14] and INTERNIST-II [15]. In HEARSAY-II. knowledge sources are activated whenever they see their prerequisite data on the blackboard, i.e., the activation of knowledge sources is data-driven. In INTERNIST-II. hypotheses are also generated from the data. Signs and symptoms are input to the system and the appropriate hypotheses are activated. Therefore, data is continually being added as a result of questions asked by the system, so that new hypotheses are introduced throughout the diagnostic process. Our scheme Tor adding hypotheses differs in that it takes advantage of both successes and failures in order to determine which alternates are viable. The remainder of this section will concentrate on some of the more interesting features of the control structure. A complete description of the control structure can be found in [16].

Our system maintains a set of active hypotheses throughout the recognition process. Activation of a hypothesis involves activation of its IS-A ancestors (via the IS-A hierarchy) and of its components (via the PART-OF hierarchy). Active hypotheses are organized by their "conceptual adjacency": the semantic inter-hypothesis relationships that are active at any time instant, thus conceptual adjacencies are dynamic. This may he likened to the "structural adjacency" of HEARSAY-11, which is based on immediate adjacency in the AND/OR tree of hypotheses. These relationships correspond to the features of the representation - IS-A, PART-OF, similarity, and temporal precedence.

An object's set of possible hypotheses consists of active hypotheses which define motion classes for that object during a common time interval. Hypotheses for an object's parts are considered as part of the set for that particular part. In this way, since hypotheses are organized by object, there can be no overlapping sets of hypotheses - no one hypothesis can be a member of more than one set. There is no restriction, however, on the PART-OF relations between sets of hypotheses.

The mechanism for the addition of a hypothesis via an activated similarity link requires elaboration. A similarity link in hypothesis H is activated between hypotheses H and H" when:

• the similarity expression is satisfied, i.e., all properties of H' which must be true before H* can be activated must have been instantiated; and,

• at least one of the exceptions in the difference expression has been instantiated. It is not necessnry that this be the first one in the time course expression. Noise effects may mask it. One must realize, however, that noise or extraneous information may also erroneously trigger the similarity.

When frame H* is activated by H as its competitor, the slots of H' must be filled up to the time instant of activation. It will be true, in many cases, that some of the slots of H will be identical in content to those of H' These are transferred through the binding expression in the similarity link. Those that are not must be computed.

There are several additional considerations that arise when determining similarities between hypotheses:

• When a particular exception cannot be handled by the local similarity links, the links of its IS-A ancestors are checked.

• Since activation of a frame automatically activates its IS-A ancestors, transfer of parts between the two frames* IS-A ancestors may be accomplished by the binding expressions of the similarity links (if present) between those frames.

• There is also a need to propagate similarities upwards along the PART-OF hierarchy, berause possible mis-matches of a motion component may require completely new contexts for the newly activated parts. This is done by automatically raising an exception in the parent hypothesis. When an exception is detected in n part A of frame B, this automatically generates an additional exception for frame B stating that frame B has failed. In this way exceptions are propagated up the PART-OF hierarchy. The exception carries with it a special slot "prereq_part" which specifies the newly activated frame name C. The parent frame of B, that is D, can use its similarity links with the added constraint that C must be PART-OF any newly activated destination frame of D.

Hypothesis Rating

The hypotheses will be ordered by means of the certainty that the system has in them. The best candidates are determined for each object; that is, hypothesis certainty values are not compared unless the hypotheses define motions for the same object and for the same time interval. Thus, each object would have a leading hypothesis set, and each part of each object would have a leading hypothesis set. It is not necessarily true that the best hypothesis for an object's part is PART-OF related to the best hypothesis for the object as a whole.

Each hypothesis has a certainty factor associated with it - a number between 0.0 (completely uncertain) and 1.0 (completely certain). Initial values are set through the conceptual adjacencies of each hypothesis when it is activated. If a hypothesis H, has N competitors (either mutually exclusive ones, through similarity links, or through the temporal precedence constraint), then each hypothesis of the competing set has an initial certainty of 1/N • cert(H). If a hypothesis has no competitors, then as far as the system is concerned, it is certain that the hypothesis will be instantiated - because there is no other possibility. This is why similarity links are so crucial - matching failures would not be recorded in a hypothesis' certainty factor through the relaxation process unless they also activate a similarity link to a competing frame. The reason for this is that certainty factors are normalized over the set of competing hypotheses. This has necessitated the inclusion of *control hypotheses:* hypotheses that monitor each sot of competing hypotheses. They are frames in the same sense as described above, but their matching constraints succeed if no hypothesis in the competing set succeeds matching and fail otherwise. They may be likened to the "no-edge" label in edge finding applications of RLP's. and, in direct analogy, are frames of type NO AGGREGATE or NO SPECIALIZATION for the relevant motion concept. Hypotheses are instantiated when their certainty reaches a particular value, determined by the size of the competing set, the duration of the shortest motion to be recognized and the image sampling rate (see [16], [17]). The instances are related to their definitional frames via the *instance of* relationship and must conform to the internal structure of the frame. Hypotheses are deleted when their certainty falls to below a minimum value that is related to the instantiating certainty threshold.

The driving concept used for the updating of certainty factors in our system is generic good temporal continuation of hypotheses. This is direct analogy to good edge continuation for edge labels in images. Global constraints extracted from the knowledge organization and local constraints embodied in the hypotheses themselves provide the basis for the determination of good temporal continuation.

Once the initial certainties are set for a hypothesis, the certainty is updated for each subsequent image. The updating rule that we use is taken from relaxation labelling [II] and has exactly the same form as presented in that paper. The compatibilities between hypotheses depend directly on the types of conceptual adjacencies that hypotheses exhibit. The weighting factor assigned to each contributing hypothesis depends on the number of active adjacent hypotheses and their types.

The major change from the standard RLP described by Zucker is that we do not iterate until certainty factors converge to stable values. Rather, certainties are updated using only one application of the rule. In this way, each update corresponds to an analyzed image in the sequence. The reason for this change requires a clear understanding of what the RLP does.

For the purposes of this argument, let us assume that there are three competing hypotheses. The initial assignment of certainty factors place the starting point of the RLP at some point, say A, on the plane formed by the certainty factor vectors fl.O.OJ, (0,1,0), }0,0, lj. The plane is as it is because the sum of the certainty factors must remain at one; otherwise, the process is not guaranteed to converge. Now suppose that after the first inter-image description is obtained, from point A. the RLP. using good temporal continuity criteria takes the certainty factor vector through some path to a final convergent labelling after several iterations of the updating rule. Why is this undesirable? Because of the nature of the RLP, the system could never move away from this final stable state, and therefore after the first pair of images, the certainty factors could not change.

If only one application of the updating rule is used between images, the effect is that the system moves along a short vector tangent to the path that it would have taken if many iterations were done. Two images are enough to determine this vector. The system's focus thus has *inertia* during the integration of many image pairs of data from the image sequence. It has inertia of location because it does not move very far away from its current location. This is desirable because hypotheses and their compatibility values are uncertain. The length of the vector is determined by how consistent the matchings of hypotheses and the kineses are with the constraints imposed by the organizational axes of the KB. The system also has inertia of direction because it moves along a tangent to its current path in the plane formed by the hypothesis space. The problem of rapidly shifting foci of attention, present in other large application systems [14], [15|, is therefore greatly reduced. However, other considerations arise. What arc the limitations of such a process in terms of how many iterations (in other words, images) are needed in order to determine a consistent labelling, and how is such a process affected by extraneous stimuli? We have conducted several experiments on the hypothesis rating scheme in order to study these questions [16], [17]. The result?; of these experiments have lead to several empirical relationships relating compatibility values for the conceptual adjacency types, maximum number of hypotheses in a single competing set, decision certainty thresholds, signal noise content, scene sampling rate, and shortest event in the domain to be recognized. In addition, the results of one particular experiment demonstrate that inclusion of the 1S-A hierarchy and its inherent generali7ation of consLraints, leads to faster event recognition in terms of the number of image samples that must be analyzed.

THE APPLICATION TO LV DYNAMICS

In surveying the rapidly growing literature on quantitative left ventricular dynamics, the problem of consistency in terminology has become apparent to us as well as to other researchers [18]. How can we compare and evaluate the varying and sometimes conflicting results of studies of ventricular wall motion? We have attempted to define the usual terms used by cardiologists in describing LV performance? in terms of the motion semantics outlined above and using our representation formalism. A brief diversion is appropriate here in order to describe the problem domain so that the example presented will be easier to digest.

The image sequences that are being studied arc obtained by cineradiography of the left ventricle. However, the particular LVs that we are studying are those that have had corrective surgery (coronary bypass or valve prosthesis). The goal is to determine how normal their post-operative function is in order to appraise the effectiveness of surgery. For each LV, during corrective surgery, 9 tiny tantalum helixes (markers) are implanted into the left ventricular wall, all roughly in the same plane. In addition, two clips are attached to the aorta as points of reference. After surgery, follow-up examination involves relatively simple cineradiography Figure 1 displays the motion patterns for one such LV, both the inward (la) and the outward (lb) motions, with all features of interest labelled. The marker positions appear as dots. The left ventricle is standardly divided into three segments, the anterior segment is made up of markers 3. *4,* 5. and 6; the apical segment is made up of markers 6. 7, and 9: and. the posterior segment is made up of markers 8. 9. 10. and 11. The aortic clips are markers 1- and *2.* These figure*? show the superimposed marker positions for a number of consecutive images, separated Into the contraction and expansion phases so that the motion paths arc more apparent Even still, some marker positions are occluded in these pictures whenever no motion was exhibited. This particular case was filmed at 30 images/sec; however, more recently. 60 images/sec has become the norm and ALVEN has also been tested on those sequences.

The left ventricular cycle has two main phases - systole (ejection, or, contraction) and diastole (filling, or, expansion). The systolic phase is divided into three parts: the isovolumic contraction phase, the maximum ejection phase and the reduced ejection phase, while diastole is divided into four phases: isovolumic relaxation, rapid filling, diastasis and filling by atrial contraction. Good data is available on the standard normal characteristics of tnese events. Using the textbook definitions of each of these phases in terms of minimum start time, maximum end time, minimum duration, maximum duration, minimum and maximum rates of contraction (expansion), frames were defined for each normal phase. Each constraint of the slots in these frames has its associated exception: TOO SHORT. TOO FAST CONTRACTION. TOO LATE. TOO EARLY, etc. TOO SHORT/LONG refers to the duration *of* an event. TOO LATE/EARLY refers to the time slot in which the motion was recognized with respect to the LV cycle. TOO FAST/b'LOW refers to the rate of area change. In addition, several other abnormalites are defined. ASYNCR0NY is the term used to denote that the onset of contraction (expansion) is not uniform for all portions of the LV. In our case, the LV is made up of three segments: anterior, apical and posterior. Each segment in turn is made up of several markers as in Figure l. Asynchrony, then, applies to each segment, for its markers, and for the LV with respect to its segments, reflecting the PART-0F hierarchy. HYP0KINESIS is the term used to describe a significantly smaller extent of motion in comparison with the object's neighbours and again is defined relative to the PART-0F hierarchy. Finally, abnormal directions of motion with respect to the contraction (expansion) of the LV are described by the term DYSKINESIS. Other motion terms used in the description aro: UNIFORM contraction (expansion) * ail parts are moving inwards (outwards) synchronously, and the superpart is contracting (expanding) (these are examples of aggregation of simultaneous motions); INWARDS

<UUTWARDS) motion with respect to the dynamic centroid of an object's super-part; length changes, both SHORTENING and LENGTHENING (normal I.Vs' perminter shortens only during contraction and lengthens only during expansion); and. other? that have a more intuitive meaning. A BEAT is defined as a strict contraction immediately followed by a strict expansion, while an INVERSE BEAT is the exact opposite (these are examples of an aggregation of a sequence of motions). Finally, the terms MOVING, TRANSLATE, OBJECTIVE. PHYSICAL-CHANGE. AKEA-CHAMGE. LENGTH-CHANGE reflect the 1S-A hierarchy for each of their more specialized motions* Also note that terms such as NO MOTION, NO TRANSLATION. NO LENGTH CHANGE, etc. are also recognized. Reflexive verbs, although defined in ALVEN. such as UPWARDS. LEFTWARDS, etc. were not included in this desciption. They are not very useful for LV motion since an orientation-free description is required. The term R-TS refers to an instance of a normal rate and time slot for a particular area change. The set of such normal instances is then put together in determining whether or not the duration is proper.

An area for the LV is defined as the area enclosed by the outline formed by joining the markers, and an area for a segment is defined as the area enclosed by the outline formed by the markers of the segment and thu dynamic ccntroid of the LV. A length is simply the arclength of the line formed by joining the markers of the LV as a whole or of each segment.

ALVEN's classification of the motions of each segment, for the LV as a whole, as well as all exceptions to normal motion generated for all markers and the complete descriptions of markers 2 and 7. makes up Figure 2. The goal is to provide a summary description of the dynamics of the LV as a whole, and then, if the clinician is interested, the segments' dynamics, marker dynamics and their associated numericn) quantities may be selectively examined for the purpose of either verification of the summary description or for precisely locating reasons for an abnormal summary description. This particular example required 17 images for a single LV cycle at 30 images/sec. The radiologist reported that the marker motions indicated a hypokinetic anterior segment and normal motions otherwise.

The description in Figure 2 uses the time units of images. Wc see by the description, That ALVEN agrees with the radiologist in that hypokinesis is reported for the contraction phase of each anterior marker for at least one time instant during the phase, as well as for the anterior segment as a whole. However, hypokinesis is just a sign of more severe problems. Obviously, if a part of the LV does not contract properly, the rate of expulsion of blood is impaired. This is apparent by the exceptions TOO SLOW CONTRACTION, TOO SHORT REDUCED EJECTION PHASE, the lack of an instance of the maximum ejection phase, etc. These are consistent with hypokinesis, but are more serious problems with the left ventricle. Also, since the system has objective criteria for judging the motions, several instances of agynchrony and dyskinesis were found both at the segment and at the LV level This is rather difficult if not impossible to do by observation only. The remainder of the description gives the complete set of terms used by ALVEN to describe the motions.

We conclude that ALVEN's evaluation is consistent with that of a radiologist, but is more complete and objective. Other examples that have been analyzed also support this conclusion.

CONCLUSIONS

We have briefly described an approach to the abstraction of time-varying concepts from a sequence of images. The system embodying this approach, called ALVEN, has been implemented and an example of one of the left ventricular motion sequences was presented. Testing thus far has shown that ALVEN's evaluation of LV performance is consistent with the radiology reports, but is more objective and complete. ALVEN has also been tested on arbitrary motions of moving dots and produces quite acceptable, although limited in vocabulary, descriptions of their motions in the plane of the image. We believe that the recognition strategy is unique in that the semantic relationships of the KB and time itself are integral components of the recogntion process. Moreover, our recognition control structure provides an approach for reducing the effects of shifting foci during recognition and for dealing with exceptions and temporal constraints. There are several ongoing research topics being pursued using this general framework: interpretation of arrhythmias from electrocardiograms, interpretation of aortic valve function in the fetus using ultrasound signals, and the representation of causality and its recognition from image sequences. In addition, ALVEN's knowledge is being refined and tested in order to achieve clinical acceptance for its results.

References

1. Ullman. S., *The Interpretation of Structure from Motion.* MIT Press, 1979.
2. Rashid, R.. "Towards a System for the Interpretation of Moving Light Displays', *IEEE Trans, on Pattern Analysis and Machine Intelligence,* Nov. 1980.
3. Flinchbaugh, b\. Chandrasekaran. b\. "A Theory of Spatio-Temporal Aggregation for Vision", OSU-CISRC-TR-BO-5. Ohio State University. 1980.
4. Nagel, H. -H., "Image Sequence Analysis: What Can We Learn from Applications", University of Hamburg, IfI-HH-M-79/80, 1979.
5. Badler, N.. Aggarwal, J. (eds). "Workshop on Computer Analysis of Time-Varying Imagery", University of Pennsylvania, 1979.
8. Lcvcsquc, H., Mylopoulos. M., "A Procedural Semantics for Semantic Networks", in *Associative Networks,* ed. by Findler, Academic Press, 1979.
7. Minsky. M., "A Framework for Representing Knowledge", in *The Psychology of Computer Vision,* ed. by Winston, McGraw-Hill, 1975.
8. Miller. G., "English Verbs of Motion: A Case Study in Semantics and Lexical memory", in *Coding Processes in Human Memory,* ed. by Martin and Melton, Winston and Sons. 1972.
9. Badler, N., "Temporal Scene Analysis: Conceptual Descriptions of Object Movements", TR-80, Dept. of Computer Science. University of Toronto. 1975.
10. Tsotsos, J.K., "A Framework for Visual Motion Understanding", CSRG TR-114, University of Toronto, I960.
11. Zucker. S.. "Production Systems with Feedback", in *Pattern-Directed Inference Systems,* ed. by Waterman and Hayes-Roth, Academic Press, 1978.

12. O'Rourke, J., Badler, N., "Model-Based Image Analysis of Human Motion Using Constraint Propagation", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Nov. 1980.
13. Gerbrands, J., Booman, F., Reiber, J., "Computer Analysis of Moving Radiopaque Markers from X-Ray Films", in *Computer Graphics and Image Processing*, Vol.11, 1979.
14. Lesser, V., Erman, L., "A Retrospective View of the HEARSAY-II Architecture", Proc. IJCAI77, Cambridge, 1977.
15. Pople, H., "The Formation of Composite Hypotheses in Diagnostic Problem Solving: An Exercise in Synthetic Reasoning", Proc. IJCAI77, Cambridge, 1977.
16. Tsotsos, J.K., Mylopoulos, J., Covvey, H.D., Zucker, S.W., "A Framework For Visual Motion Understanding" *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Nov. 1980.
17. Tsotsos, J.K., "On Classifying Time-Varying Events", Proceedings of Pattern Recognition and Image Processing Conference, Dallas, 1981
18. Daughters, G., et al., "Methods for Ventricular Wall Motion Assessment: Towards a Uniform Terminology" Proc. IEEE Computers in Cardiology, Geneva, 1979.

```
marker 1 :

marker 2 :
    NO MOTION  - time 0 to 2
    MOVING  - time 2 to 3
    SIMPLE  - time 2 to 3
    TRANSLATING  - time 2 to 3
    OBJECTIVE  with respect to marker 1 - time 2 to 3
    APPROACHING  with respect to marker 1 - time 2 to 3
    NO MOTION  - time 3 to 7
    MOVING  - time 7 to 8
    SIMPLE  - time 7 to 8
    TRANSLATING  - time 7 to 8
    OBJECTIVE  with respect to marker 1 - time 7 to 8
    APPROACHING  with respect to marker 1 - time 7 to 8
    NO MOTION  - time 8 to 16

marker 3 :
    CONTRACTION HYPOKINESIS  - time 4 to 6

marker 4 :
    CONTRACTION HYPOKINESIS   - time 3
    CONTRACTION HYPOKINESIS   - time 6

marker 5 :
    CONTRACTION HYPOKINESIS   - time 3

marker 6 :
    CONTRACTION HYPOKINESIS   - time 3

marker 7 :
    MOVING  - time 0 to 6
    SIMPLE  - time 0 to 6
    TRANSLATING  - time 0 to 6
    OBJECTIVE  with respect to Apical Segment - time 0 to 6
    MOVING OUTWARDS  with respect to Apical Segment - time 0 to 1
    MOVING INWARDS  with respect to Apical Segment - time 1 to 6
    NO MOTION  - time 6 to 7
    MOVING  - time 7 to 10
    SIMPLE  - time 7 to 10
    TRANSLATING  - time 7 to 10
    OBJECTIVE  with respect to Apical Segment - time 7 to 10
    MOVING INWARDS  with respect to Apical Segment - time 7 to 8
    MOVING OUTWARDS  with respect to Apical Segment - time 8 to 10
    NO MOTION  - time 10 to 11
    MOVING  - time 11 to 13
    SIMPLE  - time 11 to 13
    TRANSLATING  - time 11 to 13
    OBJECTIVE  with respect to Apical Segment - time 11 to 13
    MOVING OUTWARDS  with respect to Apical Segment - time 11 to 13
    NO MOTION  - time 13 to 14
    MOVING  - time 14 to 16
    SIMPLE  - time 14 to 16
    TRANSLATING  - time 14 to 16
    OBJECTIVE  with respect to Apical Segment - time 14 to 16
    MOVING OUTWARDS  with respect to Apical Segment - time 14 to 15
    MOVING INWARDS  with respect to Apical Segment - time 15 to 16

marker 8 :

marker 9 :

marker 10 :

marker 11 :
```

Figure 1a

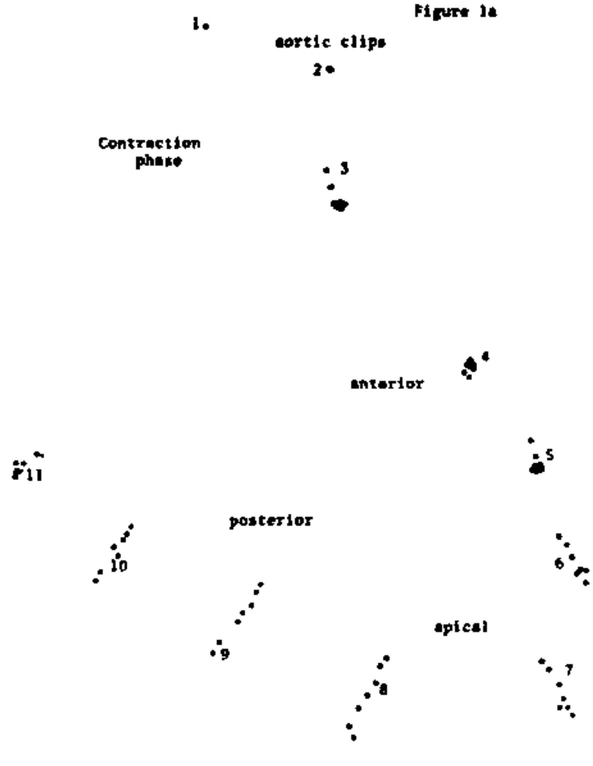
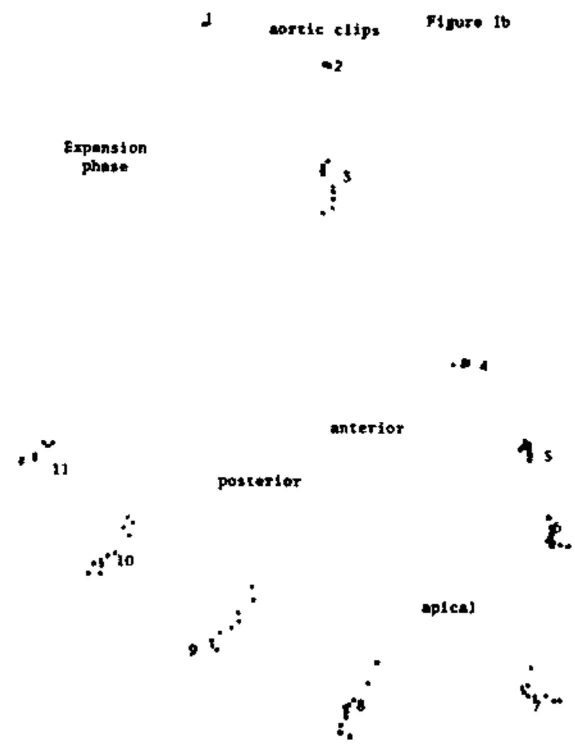


Figure 1b



```
Anterior Segment :
    CONTRACTION ASYNCHRONY   - time 1
    MOVING  - time 0 to 16
    SIMPLE  - time 0 to 16
    PHYSICAL CHANGE  - time 0 to 16
    TRANSLATING  - time 0 to 16
    AREA CHANGE  - time 0 to 16
    LENGTH CHANGE  - time 0 to 16
    OBJECTIVE  with respect to Left Ventricle - time 0 to 1
    CONTRACTING  - time 0 to 6
    SHORTENING  - time 0 to 2
    MOVING INWARDS  with respect to Left Ventricle - time 0 to 1
    CONTRACTION DYSKINESIS   - time 2
    LENGTHENING  - time 2 to 3
    OBJECTIVE  with respect to Left Ventricle - time 3 to 11
    SHORTENING  - time 3 to 6
    CONTRACTION HYPOKINESIS  - time 3 to 6
    MOVING INWARDS  with respect to Left Ventricle - time 4 to 6
    UNIFORMLY CONTRACTING  - time 3 to 5
    CONTRACTION DYSKINESIS   - time 6
    EXPANSION ASYNCHRONY   - time 7
    MOVING OUTWARDS  with respect to Left Ventricle - time 6 to 7
    EXPANDING  - time 6 to 16
    LENGTHENING  - time 6 to 12
    BEATING  - time 0 to 16
    EXPANSION DYSKINESIS  - time 8 to 16
    MOVING INWARDS  with respect to Left Ventricle - time 7 to 8
    EXPANSION HYPOKINESIS   - time 9
    MOVING OUTWARDS  with respect to Left Ventricle - time 8 to 11
    OBJECTIVE  with respect to Left Ventricle - time 12 to 16
    SHORTENING  - time 12 to 13
    MOVING OUTWARDS  with respect to Left Ventricle - time 12 to 14
    LENGTHENING  - time 13 to 14
    CONTRACTION DYSKINESIS   - time 15
    CONTRACTING  - time 14 to 15
    INVERSE BEATING  - time 6 to 15
    SHORTENING  - time 14 to 15
    MOVING INWARDS  with respect to Left Ventricle - time 14 to 15
    EXPANDING  - time 15 to 16
    LENGTHENING  - time 15 to 16
```

```
Posterior Segment :
    CONTRACTION ASYNCHRONY   - time 1
    MOVING  - time 0 to 16
    SIMPLE  - time 0 to 16
    PHYSICAL CHANGE  - time 0 to 16
    TRANSLATING  - time 0 to 6
    AREA CHANGE  - time 0 to 16
    LENGTH CHANGE  - time 0 to 6
    OBJECTIVE  with respect to Left Ventricle - time 0 to 6
    CONTRACTING  - time 0 to 6
    SHORTENING  - time 0 to 2
    MOVING INWARDS  with respect to Left Ventricle - time 0 to 6
    CONTRACTION DYSKINESIS  - time 2 to 6
    LENGTHENING  - time 2 to 3
    SHORTENING  - time 3 to 6
    CONTRACTION HYPOKINESIS   - time 6
    EXPANSION ASYNCHRONY   - time 7
    NO TRANSLATION  - time 6 to 7
    NO LENGTH CHANGE  - time 6 to 7
    EXPANDING  - time 6 to 14
    BEATING  - time 0 to 16
    EXPANSION DYSKINESIS  - time 8 to 16
    TRANSLATING  - time 7 to 16
    LENGTH CHANGE  - time 7 to 16
    LENGTHENING  - time 7 to 8
    EXPANSION HYPOKINESIS   - time 9
    OBJECTIVE  with respect to Left Ventricle - time 8 to 16
    SHORTENING  - time 8 to 9
    MOVING OUTWARDS  with respect to Left Ventricle - time 8 to 13
    LENGTHENING  - time 9 to 13
    SHORTENING  - time 13 to 14
    MOVING INWARDS  with respect to Left Ventricle - time 13 to 15
    CONTRACTION DYSKINESIS   - time 15
    CONTRACTING  - time 14 to 15
    INVERSE BEATING  - time 6 to 15
    LENGTHENING  - time 14 to 16
    MOVING OUTWARDS  with respect to Left Ventricle - time 15 to 16
    EXPANDING  - time 15 to 16
```

```
Apical Segment :
    CONTRACTION ASYNCHRONY   - time 1
    MOVING  - time 0 to 16
    SIMPLE  - time 0 to 16
    PHYSICAL CHANGE  - time 0 to 16
    TRANSLATING  - time 0 to 6
    AREA CHANGE  - time 0 to 16
    LENGTH CHANGE  - time 0 to 6
    CONTRACTING  - time 0 to 6
    LENGTHENING  - time 0 to 1
    CONTRACTION DYSKINESIS  - time 3 to 6
    OBJECTIVE  with respect to Left Ventricle - time 1 to 6
    SHORTENING  - time 1 to 6
    MOVING INWARDS  with respect to Left Ventricle - time 1 to 6
    EXPANSION ASYNCHRONY   - time 7
    NO TRANSLATION  - time 6 to 7
    NO LENGTH CHANGE  - time 6 to 7
    EXPANDING  - time 6 to 16
    BEATING  - time 0 to 16
    EXPANSION DYSKINESIS  - time 8 to 16
    TRANSLATING  - time 7 to 16
    LENGTH CHANGE  - time 7 to 16
    OBJECTIVE  with respect to Left Ventricle - time 7 to 10
    LENGTHENING  - time 7 to 13
    MOVING OUTWARDS  with respect to Left Ventricle - time 7 to 9
    MOVING INWARDS  with respect to Left Ventricle - time 9 to 10
    OBJECTIVE  with respect to Left Ventricle - time 11 to 16
    MOVING OUTWARDS  with respect to Left Ventricle - time 11 to 16
    SHORTENING  - time 13 to 14
    CONTRACTION DYSKINESIS   - time 15
    CONTRACTING  - time 14 to 16
    INVERSE BEATING  - time 6 to 16
    LENGTHENING  - time 14 to 16
    OBJECTIVE  with respect to Left Ventricle - time 15 to 16
    MOVING OUTWARDS  with respect to Left Ventricle - time 15 to 16
    EXPANDING  - time 15 to 16
```

```
Left Ventricle :
    CONTRACTION ASYNCHRONY   - time 1
    MOVING  - time 0 to 16
    SIMPLE  - time 0 to 16
    PHYSICAL CHANGE  - time 0 to 16
    TRANSLATING  - time 0 to 16
    AREA CHANGE  - time 0 to 16
    LENGTH CHANGE  - time 0 to 16
    CONTRACTING  - time 0 to 6
    SHORTENING  - time 0 to 6
    TOO SLOW CONTRACTION   - time 1
    CONTRACTION DYSKINESIS  - time 3 to 4
    UNIFORMLY CONTRACTING  - time 4 to 6
    NORMAL R-TS REDUCED EJECTION  - time 3 to 6
    TOO SHORT REDUCED EJECTION PHASE   - time 7
    EXPANSION ASYNCHRONY   - time 7
    EXPANDING  - time 6 to 14
    LENGTHENING  - time 6 to 16
    TOO SLOW EXPANSION FOR RAPID FILL  - time 7 to 10
    BEATING  - time 0 to 16
    EXPANSION DYSKINESIS  - time 8 to 16
    UNIFORMLY EXPANDING  - time 7 to 10
    TOO FAST EXPANSION FOR DIASTASIS   - time 11
    NORMAL R-TS DIASTASIS  - time 11 to 12
    TOO SHORT DIASTASIS  - time 13 to 15
    NORMAL R-TS DIASTASIS  - time 13 to 14
    CONTRACTION DYSKINESIS   - time 15
    CONTRACTING  - time 14 to 15
    INVERSE BEATING  - time 6 to 15
    NORMAL R-TS ATRIAL FILLING  - time 15 to 16
    EXPANDING  - time 15 to 16
```