ALVEN: A SYSTEM FOR LV SEGMENTAL WALL MOTION ANALYSIS*

J.K. Tsotsos, H.D. Covvey, J. Mylopoulos, S. Zucker, and E.D. Wigle

Cardiovascular Unit(UC-1-104), Toronto General Hospital
101 College Street
Toronto, Ontario, Canada

This paper describes a framework for the analysis and classification of human left ventricular (LV) segmental wall motion, realized in a system called ALVEN.

ALVEN has two major components: a data collection module and an analysis and description module. The data collection module is used for gathering statistics on LV wall motion, both normal and abnormal. Statistics are computed from the sequences of outlines, providing data later used in the analysis and description module of ALVEN.

Introduction

The detailed analysis of cardiac images from both invasive and non-invasive studies continues to be a burdensome and costly problem. Despite over 15 years of research and development a truly successful automatic ventricular image analysis system does not yet exist. We have made a new approach to the basic problem of ventricular border recognition and, have made our target a complete ventricular wall motion analysis and description system.

Our research has impact in two major research areas -- artificial intelligence and medical science. Our objectives are: the advancement of design methodologies for computer systems that analyze motion, and the development of a computer system that can perform analysis of left ventricular wall motion by processing cinecardioangiograms. Such a system can serve as a formal model of the mechanics of left ventricular wall motion and also as an assistant to physicians who wish to make the process of analysis more consistent and objective and less time consuming. In addition, hypotheses about left ventricular dynamics are only now evolving, and a system such as ours could be used as a testbed for such hypotheses.

Our thesis is that a system whose control structure is based on competition and cooperation among hypotheses, and that is driven by a knowledge base (KB), is adequate for classification of a restricted set of motions exhibited in an image sequence. We will call such a system a Motion Understanding System (MUS). In addition, we claim that the representational formalism we have defined for the knowledge base is sufficient for representing the motion concepts necessary for general motion understanding, and in particular, for left ventricular wall motion. The formalism, however, is restricted in that it cannot represent causal relationships between motions, and it only deals with two-dimensional motions and therefore occlusion is not handled.

*WORK SUPPORTED BY THE ONTARIO HEART FOUNDATION

The KB must contain both general and problem-specific motion concepts. The recognition algorithms use this KB in order to assist them in determining which motion classes are exhibited by objects in an image sequence. Several assumptions are implicit in the recognition control structure:

1. The motions in the image sequence are continuous, i.e., there are no significant motions occurring between two images.
2. Each motion exhibited occurs over several images (perhaps necessitating a high scene sampling rate).
3. As in other knowledge based understanding systems, we assume that the KB has predictive power, i.e., it can be used in order to predict object locations in each image, and possible motion classes exhibited by each object.

Methodology

It is important to define the limits that we impose on our methodology. We assume that such a system will be used to do question-answering about the films it is presented. Questions are asked before the sequence is analyzed and replies come only after the entire sequence has been analyzed. The reply will be either at the same level of abstraction or at a more general one -- never at a more specific one.

Our methodology integrates several current Artificial Intelligence techniques with several novel ones. A representation formalism for general temporal events has been designed that is intended to handle all of the semantic components deemed necessary for adequate motion description in two dimensions. Knowledge units are defined for each motion concept and are organized using concepts from semantic network theory, particularly the PSN formalism.[1] An intermediate language, using this formalism, is provided for linking quantitative image data with qualitative conceptual information. Using our representation formalism, a KB of general motion concepts is defined. In addition, a problem-specific KB for left ventricular wall motion is presented which uses general motion concepts for motion description.

The intermediate language that we use for "inter-image descriptions" is made up of five primitive temporal concepts: TIME-INTERVAL, LOCATION-CHANGE, AREA-CHANGE, LENGTH-CHANGE, and

SHAPE-CHANGE. Each of these concepts has associated components that define their semantics. For example, TIME-INTERVAL has a "start-time," and "end-time" and a "duration," while LOCATION-CHANGE has a "start-point," and "end-point" and a "time-interval" over which the motion occurred.

Each knowledge unit in the KB has semantic components, thus defining a PART-OF hierarchy. In addition, knowledge units are organized into an IS-A hierarchy that relates levels of abstraction.

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 KB, that is, by the primitive relations between knowledge units. A feedback loop is incorporated in order to link the several components of the structure. The low level vision process is expectation guided, with modifications to the relaxation labelling process[2] integrating high level expectations with edge determination. High level expectations combine knowledge about expected extents of motion for a given time interval with the history of the actual motions observed. The low level process uses a prediction window along with preferred edge orientations in order to bias the relaxation process.

Detected motions and inter-image descriptions of motions are described using the set of primitive motion concepts that comprise the intermediate description language. These inter-image descriptions are matched against hypothesized motions and differences are detected. Similarity links provide the basis of the change of attention mechanism that acts as an exception handling mechanism. These links, however, do not join two knowledge units unless events in the image sequence deem it necessary. With each potential link is stored a time course of differences between the 2 potentially linked knowledge units. When the differences specified are actually observed, the link is activated. A change of attention in our system does not mean that one hypothesis is superceeded by another: rather, both are considered as competitors. Since several simultaneous hypotheses can co-exist, a focus of attention mechanism is necessary in order to limit the number of hypotheses under consideration. Our focus of attention mechanism uses a property of feedback systems: inertia, i.e., the output changes slowly over time and therefore changes in focus are continuous. The system focuses on the best hypotheses under consideration. Certainty factors attached to hypotheses determine their ranking. The certainty factors are updated using relaxation labelling with dynamic neighbourhoods and compatibilities that are determined using the conceptual adjacency between hypotheses and matching progress. A single application of relaxation between images preserves the inertia of the feedback system.

Several important problems are directly faced by our control structure: discrimination between 2 hypothesized motion states; noise; and temporal segmentation (i.e. when does one motion end and the next begin).

First let us define the terms. We will say that the system has discriminated between 2 concurrent labels (hypotheses) if the difference in their certainty values becomes 0.667 or greater. Empirical results have shown that under ideal, noisefree conditions, the system cannot discriminate between two simultaneously hypothesized motions if the motion is exhibited in fewer than three images. In non-ideal conditions, this lower limit increases. The system can be tuned to enable it to recognize the motion of shortest duration in a given problem domain, up to this limit, by adjusting the strength of competition between hypotheses that are linked by a similarity link.

Noise may be accomodated in the process explicitly. An estimate of the amount of noise can be used to adjust the strength of cooperation between hypotheses that satisfy PART-OF relationships.

One way by which the temporal segmentation problem may be solved is to define the time instant of segmentation to be the time instant at which the certainty of one motion hypothesis exceeds that of its competitor's for that time interval. However, experiments have shown that an independent temporal segmentation process is necessary if, in the best case, the time interval spanned by 6 images presents a significant error to the determination of event durations in the problem domain. The system's error in duration determination, up to this limit, may be controlled by adjusting the strength of competition for hypotheses with overlapping expected time intervals.

Conclusions

Several portions of this system have been implemented and successfully tested. The technical details and several examples of the application to cardiac images appear in the recent doctoral thesis of the first author.[3]

References

1. Levesque, H., Mylopoulos, J., "A procedural semantics for semantic networks," In Associative Networks, Findler(ed.), Academic Press, 1979.
2. Zucker, S.W., Hummel, R.A., Rosenfeld, A., "An application of relaxation labelling to line and curve enhancement," IEEE Transactions on Computers, Vol. C-26, April 1977.
3. Tsotsos, J.K., "A framework for visual motion understanding," Ph.D. thesis, Dept. of Computer Science, University of Toronto, December 1979.