Dependence of Recognition Accuracy on Available Network Bandwidth

Michael D. DeVore
Joseph A. O'Sullivan
Roger D. Chamberlain
Mark A. Franklin

Michael D. DeVore, Joseph A. O'Sullivan, Roger D. Chamberlain, and Mark A. Franklin, "Dependence of Recognition Accuracy on Available Network Bandwidth," in *Proc. of 5th High Performance Embedded Computing Workshop*, November 2001.

Department of Electrical Engineering
Washington University
St. Louis, Missouri

Dependence of Recognition Accuracy on Available Network Bandwidth*

Michael D. DeVore[1], Joseph A. O'Sullivan[1], Roger D. Chamberlain[2], Mark A. Franklin[2]
[1]Electronic Systems and Signals Research Laboratory
[2]Computer and Communications Research Center
Department of Electrical Engineering, Washington University, St. Louis MO 63130

1 Introduction

Automatic recognition can be formulated as a search over a database of object models for the model class that is most likely to have yielded a given observation. An automatic recognition system may be implemented as a collection of processors which share an imaging sensor platform. Such a system may be one module of a larger, interconnected system with both the object model database and interconnection network operating as shared resources. Examples include a recognition process executing on a computing platform, sharing a bus and magnetic storage media with other applications, and a network of embedded recognition systems sharing an object model database.

In many cases, a recognition module can be designed so that, at any given time, the classification decision represents the most likely object class found up to that time. In this way, the accuracy of a classification depends upon the extent of the model database searched prior to making a decision. Decisions can be made incrementally, with an initial classification successively refined as the search proceeds. For such systems, it is desirable to map the recognition algorithm to the available architecture in such a way that the probability of classification error decreases rapidly at the beginning of the search.

For a given average number of observations classified per unit time, the fraction of the database that can be searched depends upon both the module's computational resources and the bandwidth available for communicating with the model database. For a given architecture and assuming that processing and communication do not completely overlap, more available bandwidth implies that more of the database can be searched in a given amount of time. Thus, the available network bandwidth has a direct impact on the accuracy of the recognition module.

We consider architectures that facilitate incremental database searches employing successively-refinable object models that encode object properties in a coarse-to-fine hierarchy. We further develop computational performance models for these recognition modules in terms of parameters describing the architecture and the available network bandwidth. The computational models are combined with empirical recognition accuracy results for actual synthetic aperture radar imagery and directly relate accuracy to the available network bandwidth as a case study.

2 Successive Refinability

Objects imaged by a sensor platform may exhibit a wide variability in pose relative to the platform. Finding the most likely explanation for a given observation involves a search over all possible objects and all poses of those objects. Presumably, the number of distinct object classes recognizable by a module is finite, however their poses are continuous-valued with six or more degrees of freedom. Object pose can be discretized, but accounting for pose variation can easily dominate the search time due to the many possible combinations of pose parameters.

Classification decisions can be successively refined by basing an initial decision on a coarse discretization of object pose and successive decisions on more finely discretized pose [1]. The average accuracy of a recognition algorithm can be represented as a function $f(\gamma; \alpha_{\mathrm{atr}})$, where γ characterizes the extent of refinement or, equivalently, the fraction of the model

*This work is supported in part by the DARPA VLSI Photonics Program under grant DAAL01-98-C-0074, by the Boeing Foundation, and by the Office of Naval Research grant N00014-98-1-06-06.

database searched and α_{atr} is a parameter vector characterizing the recognition problem. Refinable decisions are facilitated through object models that are successively refinable over pose. Such models represent objects with high pose resolution in terms of low pose resolution models with some incremental information. A recognition module which has already received a low resolution model can be sent a high resolution model by communicating only the incremental information needed. The accuracy $f(\gamma; \alpha_{\mathrm{atr}})$ can be determined empirically over a collection of sample data for a given recognition problem.

3 Network Dependence

From knowledge of the algorithm employed and the module architecture, computational models can be constructed for the rate at which the object model database is searched [2]. This rate, $R(\alpha_{\mathrm{comp}}, \mathrm{BW})$ is the number of discrete combinations of object pose and class considered per second and is a function of the computational resources described by α_{comp} and the available network bandwidth, BW. Given a computational architecture α_{comp} and time limits on observation classification, the extent of the database searched, and hence the accuracy f, depend upon the available network bandwidth.

To search through some fraction γ of the database requires that $b(\gamma)$ bits be transmitted to the recognition module and the module must spend $T_{\mathrm{proc}}(\gamma)$ seconds processing. If each observation must be classified in T_{class} seconds and there is no overlap between processing and communication, then the minimum available bandwidth must be

$$BW(\gamma) = \frac{b(\gamma)}{T_{\mathrm{class}} - T_{\mathrm{proc}}(\gamma)}.$$

Figure 1 shows plots of probability of classification error versus required bandwidth for a fixed computational architecture and several required recognition times T_{class}. For this example, the architecture parameter α_{comp} characterizes a module with four superscalar processors with a 1 GHz clock rate. The recognition problem characterized by α_{atr} is a ten-class SAR recognition problem with known radar depression angle and one-foot resolution. For each curve, increasing available bandwidth allows more of the database to be searched before a final decision is reached, and the probability of error decreases. For a given probability of error, increasing bandwidth permits faster classification. Alternatively, for a given bandwidth, shorter required classification times result in lower classification accuracy.

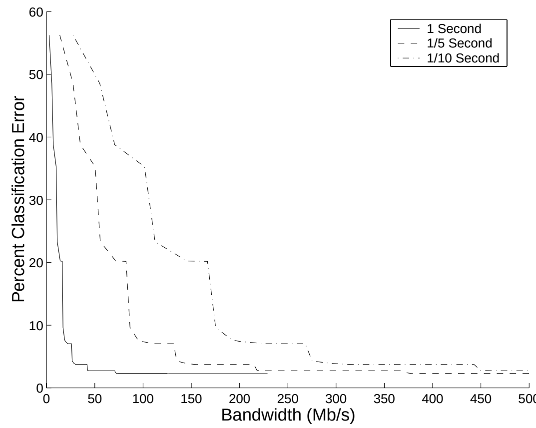

Figure 1: Probability of classification error versus average required bandwidth for several observation classification times.

Analyses of this type can be useful during the research, design, and operation of object recognition systems. The accuracy-bandwidth curve can serve as a measure of how quickly a given database search algorithm locates likely explanations for observations and can be used to compare competing algorithms. During system design, the analysis points to the network bandwidth that must be reserved for a recognition module in order to meet design goals on accuracy and throughput. Finally, for an operating unit, the accuracy-bandwidth curve characterizes the change in recognition accuracy if network bandwidth must be dynamically reallocated to higher priority processes.

References

[1] M. D. DeVore, J. A. O'Sullivan, S. Anand, and N. A. Schmid. Probabilistic approach to model extraction from training data. In E. G. Zelnio, editor, *Algorithms for Synthetic Aperture Radar Imagery VIII, Proc. of SPIE*, 2001.

[2] M. D. DeVore, J. A. O'Sullivan, R. D. Chamberlain, and M. A. Franklin. Relationships between computational system performance and recognition system performance. In F. A. Sadjadi, editor, *Automatic Target Recognition XI, Proc. of SPIE*, volume 4379, 2001.