Template for COSC 387 Research Papers

Marcus A. Maloof

Department of Computer Science
Georgetown University
Washington, DC 20057-1232

maloof@cs.georgetown.edu

1 September 1999

Abstract

We consider the problem of detecting rooftops in overhead images, which is one processing step in a building detection system. Currently, the system uses a hand-configured linear classifier to select the most promising rooftop candidates for further processing. We present results from an empirical study in which we used machine learning methods to acquire the selection criteria for rooftops. ROC analysis demonstrated that a naive Bayesian classifier performed better than the hand-configured solution, using area under the ROC curve as the measure of performance.

1 Introduction

This is my introduction.

2 Problem Statement

3 Survey of the Literature

Research on learning in computer vision has become increasingly common in recent years. Some work in visual learning takes an image-based approach (Beymer & Poggio, 1996), in which the images themselves, usually normalized or transformed in some way, are used as input to a learning process, which is responsible for forming the intermediate representations necessary to transform the pixels into a decision or classification. Researchers have used this approach extensively for face and gesture recognition (Chan, Nasrabadi, & Mirelli, 1996; Osuna, Freund, & Girosi, 1997), although it has seen other applications as well (Nayar & Poggio, 1996; Pomerleau, 1996).

A slightly different approach relies on handcrafted vision routines to extract relevant image features, based on intensity or shape properties, then recognizes objects using learned classifiers that take these features as inputs. For example, Shepherd (1983) used decision-tree induction (Quinlan, 1993) to construct classifiers for chocolate shapes in an industrial vision application. Cromwell and Kak (1991) took a similar approach to recognizing electrical components, such as transistors, resistors, and capacitors.

Several researchers have also investigated learning for three-dimensional vision systems. Papers by Conklin (1993), and Connell and Brady (1987) describe inductive approaches aimed at improving object recognition. The aim here is to learn the three-dimensional structure that characterizes an object or object class, rather than its appearance. Another line of research, which falls midway between this approach and image-based schemes, instead attempts to learn a small set of *characteristic views*, each of which can be used to recognize an object from a different perspective (Gros, 1993; Pope & Lowe, 1996).

Blah, blah, blah...

4 Experimental Study

4.1 Method

4.2 Results

4.3 Analysis

5 Conclusion

In this paper we have described GTDR, a system for generalizing temporal diagnosis rules, which is useful for data mining in manufacturing applications. As we have shown, GTDR rests upon a strong theoretical basis and extends this original theory by considering temporal variables rather than atemporal ones. Future work will involve developing routines to reduce the number of temporal rules.

Acknowledgements. The author thanks Paul Jones who helped clarify many of the ideas discussed in this report, and Jacob Fredman for providing the software for feature extraction.

References

Beymer, D., & Poggio, T. (1996). Image representations for visual learning. *Science*, *272*, 1905–1909.

Chan, L., Nasrabadi, N., & Mirelli, V. (1996). Multi-stage target recognition using modular vector quantizers and multilayer perceptrons. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 114–119). Los Alamitos, CA: IEEE Press.

Conklin, D. (1993). Transformation-invariant indexing and machine discovery for computer vision. In *Working Notes of the AAAI Fall Symposium on Machine Learning in Computer Vision* (pp. 10–14). Menlo Park, CA: AAAI Press.

Connell, J., & Brady, M. (1987). Generating and generalizing models of visual objects. *Artificial Intelligence*, *31*, 159–183.

Cromwell, R., & Kak, A. (1991). Automatic generation of object class descriptions using symbolic learning techniques. In *Proceedings of the Ninth National Conference on Artificial Intelligence* (pp. 710–717). Menlo Park, CA: AAAI Press.

Gros, P. (1993). Matching and clustering: Two steps towards automatic object model generation in computer vision. In *Working Notes of the AAAI Fall Symposium on Machine Learning in Computer Vision* (pp. 40–44). Menlo Park, CA: AAAI Press.

Nayar, S., & Poggio, T. (Eds.). (1996). *Early visual learning.* New York, NY: Oxford University Press.

Osuna, E., Freund, R., & Girosi, F. (1997). Training Support Vector Machines: an application to face detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 130–136). Los Alamitos, CA: IEEE Press.

Pomerleau, D. (1996). Neural network vision for robot driving. In S. Nayar & T. Poggio (Eds.), *Early visual learning* (pp. 161–181). New York, NY: Oxford University Press.

Pope, A., & Lowe, D. (1996). Learning probabilistic appearance models for object recognition. In S. Nayar & T. Poggio (Eds.), *Early visual learning* (pp. 67–97). New York, NY: Oxford University Press.

Quinlan, J. (1993). *C4.5: Programs for machine learning.* San Francisco, CA: Morgan Kaufmann.

Shepherd, B. (1983). An appraisal of a decision tree approach to image classification. In *Proceedings of the Eighth International Joint Conference on Artificial Intelligence* (pp. 473–475). San Francisco, CA: Morgan Kaufmann.