Experimentation and Laboratory-Oriented Studies Division (DELOS), session number 1526. ID#2002-1954

ASEE Abstract Title: **Optimized signal/image feature recognition for Machine Learning. NSF** AWARD # 9980296, CRCD: Machine Learning: A Multidisciplinary Computer Engineering Graduate Program.

Sunanda Mitra, Department of Electrical and Computer Engineering, Texas Tech University, Lubbock, Texas 79409. E-mail: Sunanda.Mitra@coe.ttu.edu.

Abstract

This paper describes some of the research projects, facilitating machine learning, completed by graduate students supported by the NSF-CRCD AWARD # 9980296 entitled "Machine Learning: A Multidisciplinary Computer Engineering Graduate Program " to Texas Tech University. The program is now under development in parallel at Texas Tech and The University of Missouri at Rolla. As part of the curriculum development, courses were taught in adaptive optimization for signal processing, optimization in information theory and coding, adaptive pattern recognition, neural networks and adaptive critics, and mathematical methods and algorithms for signal processing. Thirty-five graduate students and twelve undergraduate students were significantly involved in both the research and educational activities associated with the program. Research activities were wide-ranging, and included optimized design of lossless and lossy compression for medical images, adaptive pattern recognition, segmentation, adaptive critic designs, Q-learning, optimized blind source separation, fuzzy modeling, and vector quantization. Three specific doctoral level projects involving optimization methods in signal/image processing leading to machine learning have been chosen for this paper since these projects included additional students at Master's and senior undergraduate levels in Electrical and Computer Engineering demonstrating a successful pyramid learning structure using a top down approach. Significant collaboration with federal laboratories, industries, and other universities were developed during the design and development of the projects described in the paper.

Introduction

The design and development of three specific projects, on optimization in signal/image processing and providing significant contribution to machine learning through a pyramid learning structure from senior undergraduate to doctoral level students involved in the projects, will be presented as a part of the NSF Showcase at the ASEE2002 annual meeting. The three chosen projects are:

1. Vadim Kustov, Ph.D. Dissertation, "Adaptive Filter Banks for Digital Signal Processors", December, 2000.

2. Shuyu Yang, Ph.D. Dissertation, "Performance Analysis From Rate Distortion Theory of Wavelet Domain Vector Quantization Encoding", December, 2002.

3. Zhanyu Ge, Ph.D. Dissertation, " Automated Object Recognition by Reinforcement Learning ", May 2002.

A number of Master's degree candidates in Electrical Engineering, and B.S.E.E candidates were involved in each of the above doctoral dissertation projects.

Project description

1. **Adaptive Filter Banks for Digital Signal Processors**

Discrete wavelet transform has been used in many image/signal processing applications in recent years. However, the design of optimized and adaptive wavelet filter banks is still a significant research topic, specifically in image/signal compression. A number of wavelet-based advanced lossy compression algorithms provide high-fidelity reconstruction of input images at computationally intensive costs. The present work investigates the potential and the limitations of optimized adaptive design of two-channel perfect reconstruction filters when the signal in a channel is subjected to coarse quantization during the encoding process of such advanced compression algorithms.

A real-time optimal two-channel perfect reconstruction filter bank design algorithm has been developed and implemented in a digital signal processor. The algorithm has been used in a newly developed execution time reduction method to reduce the computational costs and data storage requirement of image compression algorithms. A reduction of execution time by two to three times has been achieved without adding appreciable distortion to the reconstructed image. A number of graduate and undergraduate students worked with the doctoral student. Additional students at all levels are currently working on extensions of this project under the supervision of the principal investigator of the NSF-CRCD award at Texas Tech University..

2. **Performance Analysis From Rate Distortion Theory of Wavelet Domain Vector Quantization Encoding**

Pattern classification problems involve constrained optimization in the form of minimization of chosen cost functions. Such embedded constraints in integrated neural-fuzzy pattern recognition systems improve the performance of these systems. The comparative performance of an ART-based pattern classifier integrated with fuzzy optimization constraints is demonstrated in designing vector quantizers.

In pattern classification applications, it is important to have unsupervised learning algorithms that are stable while able to preserve significant past learning yet remain adaptable enough to incorporate new information, or new classes, when they appear. For some data mapping cases, for example, vector quantization, pattern classification algorithms are required not only to discover new clusters but also to create the number of clusters as desired, since the larger the number of clusters created, the higher the fidelity of the quantized signal. The adaptive resonance theory (ART) is a good candidate for such cases. However, since the ART networks are better for cluster discovery instead of data mapping, additional constraints are needed to design a good vector quantizer.

Algorithms combining neural network system ART and fuzzy logic have been shown to achieve better pattern recognition performance than those of ART or fuzzy algorithms alone. An adaptive fuzzy leader clustering (AFLC) algorithm is proposed for pattern recognition. AFLC has been demonstrated to be an efficient clustering algorithm in various applications, for example, noise removal, image segmentation and image vector quantization. However, the combined structure of AFLC makes mathematical analysis of the algorithm difficult. Particularly, when we consider applying AFLC to vector quantization codebook training, the sample-order-dependency of ART has to be carefully investigated.

In this research work, the goal was to analyze the performance of AFLC and provide comparative performance of two other clustering algorithms, namely, deterministic annealing (DA), and the LBG algorithm in vector quantization application to find an optimal solution. All algorithms are designed to reduce the mean square error (MSE). AFLC adopts the ART structure to discover new classes while optimizes the centroid locations (minimizes MSE) with fuzzy C-means (FCM). DA originates from the information theory and statistical mechanics. It carries out clustering such that the average distortion is minimized depending on the probability distribution of the source while the joint entropy H(X,Y) of the source X and the centroids Y is maximized. It is also speculated that DA is capable of achieving global minima. Both AFLC and DA are constrained optimization algorithms. In LBG, no specific constraint is used. Clustering is carried out by iterative updates of the centroid locations and quantization regions in order to lower the distortion (MSE). It is well known that LBG can be easily trapped in local minima.

In the problem of designing generalized codebooks for vector quantization by clustering large amount of representative training samples, it is required that the clustering algorithms be able to discover and adapt to the source distribution and thus to yield quantizers that are optimal to the source. In our current research, we are most interested in applying clustering algorithms for image vector quantization in the wavelet domain. Research has shown that for natural images, the wavelet coefficients tend to possess a single-mode distribution similar to Gaussian. Therefore, a performance comparison of the above candidate clustering algorithms is necessary for us to select an appropriate clustering algorithm for codebook training. To provide an objective comparison independent of the selected images, simulated random vectors are used as test sources.

The results from the above experiments provide very good information on how to select the appropriate clustering algorithm for vector quantization codebook training. For image compression using vector quantization, codebooks are often trained from large amount of samples whose statistical distribution is usually unknown. Wavelet coefficients of natural images tend to distribute similar to Gaussian, however, at our current stage of research, distribution of mixed coefficients from different images are not characterized when they are used for codebook training. Therefore, LBG is chosen, as we can see that LBG gives the lowest MSE in results from experiments 1 and 2 (Figures 1 and 2).

Our experiments show that in the vector quantization application, AFLC is able to achieve distortion between those of DA and LBG. Compared with DA and LBG, AFLC classifies as data appear, it does not require the knowledge of the all samples. AFLC classifies patterns as a process of discovering dissimilar events (or even rare events), which is very typical of ART. The distortion minimization property inheriting from FCM learning helps AFLC reduce the MSE and thus keep its distortion rate curves in between of DA and LBG. In this sense, AFLC enjoys the better aspects of both ART and FCM. We also demonstrate that for vector quantization application, the sample-order dependent problem of AFLC can be solved by choosing other initial centroid selection rules without compromising its ability to discover new clusters. In fact, we showed that in the improved implementation AFC, not only the problem disappears, but also the distortion is reduced. AFC is still under development in our research laboratory.

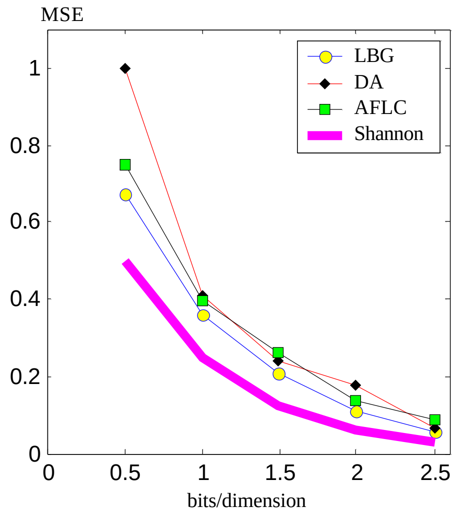


Figure 1. Quantizer distortion rate for experiments 1

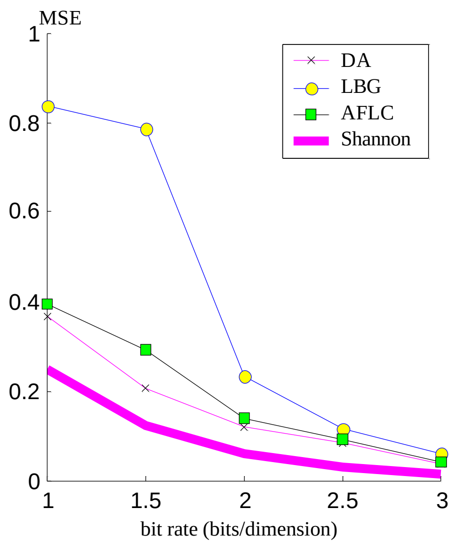


Figure 2. Quantizer distortion rate for experiments 2

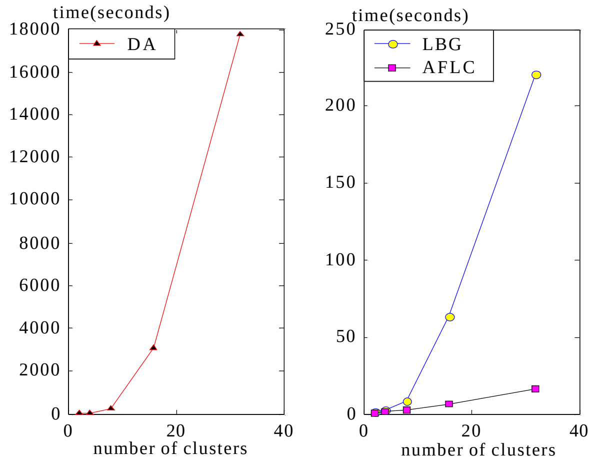


Figure 3. Clustering time for experiments 1

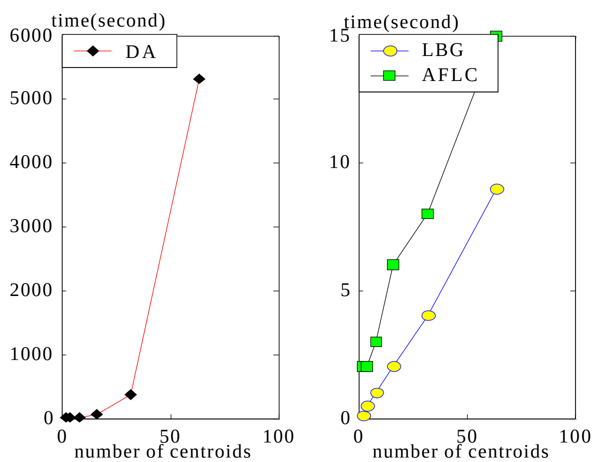


Figure 4. Clustering time for experiments 2

3. Automated Object Recognition by Reinforcement Learning

Object recognition, a branch of pattern recognition, is to identify and localize one or more objects in a given scene. We have to determine what is present and where it is within the input image. Although great achievements have been made during the last decades, currently existing object recognition techniques have shortcomings like unreliability and inefficiency, general inadaptability, manual template marking heavily influenced by human factors, inability to recognize an object without a model, and so on. Any recognition problem can be formulated as a searching process and has to be guided in a controlled manner. All search problems are optimization ones, so object recognition requires optimization and control techniques. Reinforcement learning is learning how to behave given a situation and possible actions to maximize the total expected reward in the long run. It is an optimization technique. Most pattern recognition techniques do not combine reinforcement learning for feature understanding. In the proposed work, reinforcement learning is applied both to automatic template generation from a model image and to template matching within the input image. The newly designed affine parameter estimation algorithm provides reliable results based on information contained at all feature point locations. The points are extracted in the scale-space using isophote curvature extreme points, which are invariant to affine transformations. Affine parameter estimation algorithm is applicable to any kind of translations, rotations, and scales, and decent occlusions and deformations of the object to be recognized. Preliminary results showed that the proposed set of algorithms are fast, efficient and robust. The automatic template generation algorithm, an efficient contour tracing one in gray-level images, can also be used in object recognition without a model. This is a new research field, and great amount of future work need to be done before an intelligent recognition system, as efficient as the human vision system, can be developed.

The following tasks have already been completed:

1. A complete set of algorithms for object recognition using reinforcement learning have been proposed. These algorithms are fast, efficient and robust.

2. Model template is generated from the model image without human interference.
The model template generation algorithm can also be used for object recognition without needing a prior model.

3. The algorithm for contour tracing is the first fast and efficient one for gray-level images as far as we know.

4. The affine transformation parameter estimation algorithm is fast and provides more reliable results because the algorithm uses information contained at all interest points.

5. Introduction of the energy concentration to the backward induction algorithm during template deformation contributes to a better final contour.

6. This recognition system can detect objects with any translation, rotation, scale and decent deformation and occlusion.

7. This is a "modest" system, it can learn merits from any other systems which perform better in some aspects.

This is a new research field. There is still a long way to go before a recognition system as "smart" as a human being can be developed.

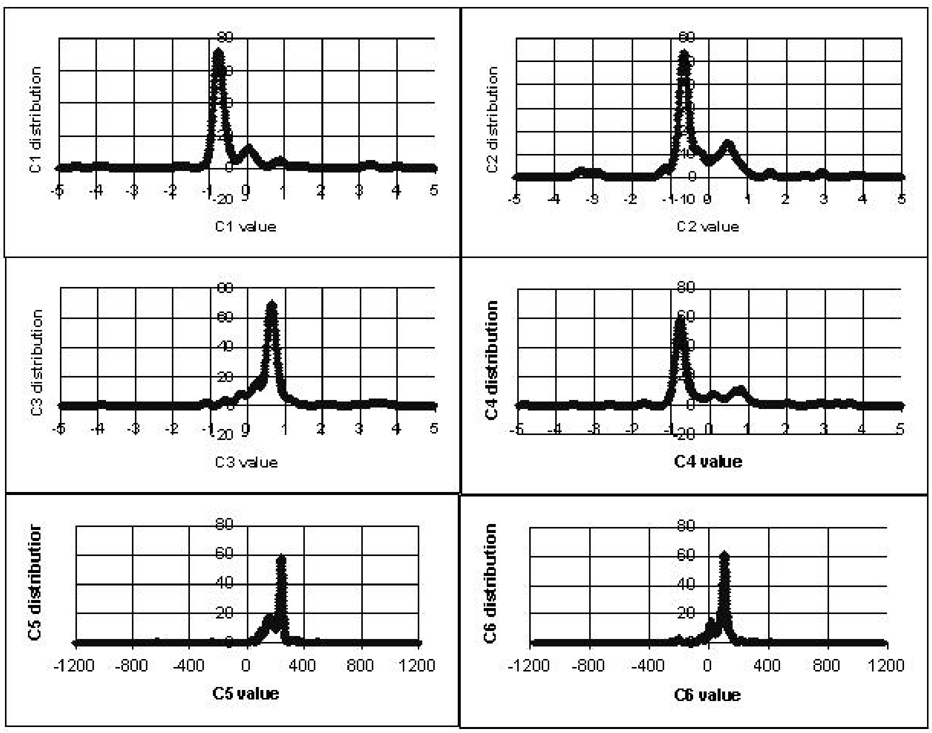


Fig. 5. Voting Curves for the Affine Transformation Parameters

The peak values of the voting curves are: C_1 = -0.74, C_2 = -0.65, C_3 = 0.65, C_4 = -0.75, C_5 = 240, C_6 = 106. These values are the estimated values of the affine transformation parameters.

An Example of Template matching in the input image

The parameters during reinforcement learning are: distance factor ? = 0.5, forgetting factor ? = 0.5, the same with those for the model reinforcement learning. The punishment factors for distance, curvature, and inhomogeneity are $\lambda_{i1} = 30, \lambda_{i2} = 10, \lambda_{i3} = 30$.

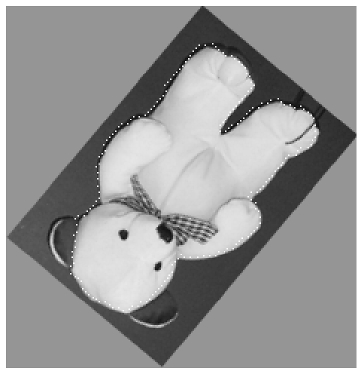
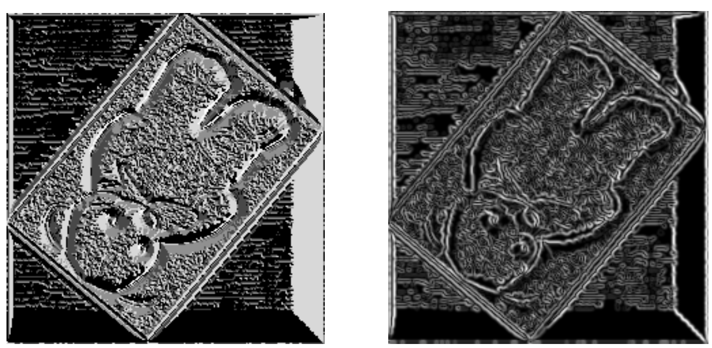
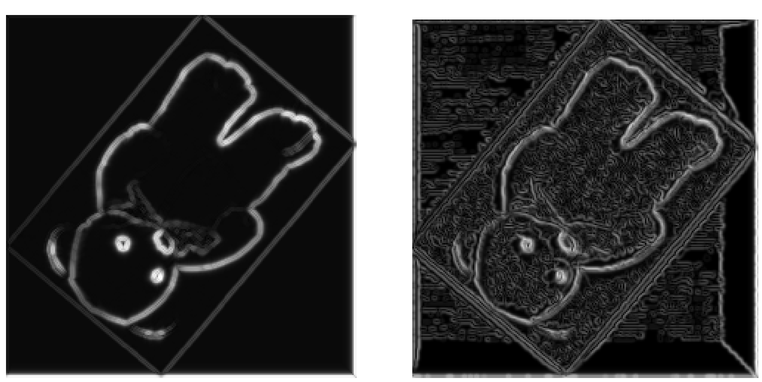


Fig.6. Template Transformed into the Input Image



Input image policy Input policy gradient

Fig. 7. Input Policy Image and Its Gradient Image



Input optimal value Input optimal value plus policy gradient

Fig. 8. Optimal Value Image and the Synthesized Image

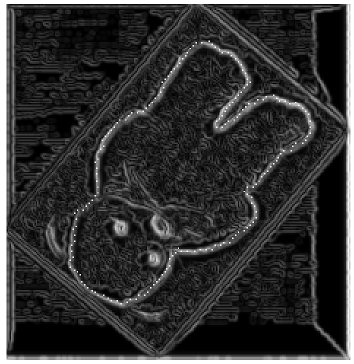


Fig. 9. Template in Synthesized Image

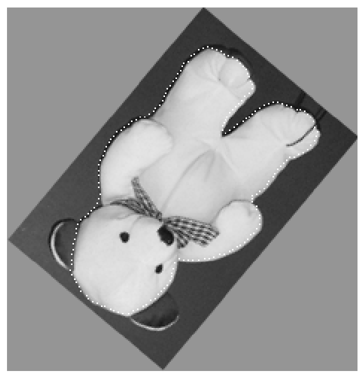


Fig. 10. Final Results after Template Deformation

Future work will address the following tasks:

(1) Apply ε-greedy learning to the learning process to make the algorithm more robust

(2) Test more model and input images, including those of changed gestures and being occluded

(3) How to choose parameters during reinforcement learning and affine parameter estimation

(4) Apply time difference learning (including Q-learning) to object recognition (contingent upon time available)

The following organizations provided collaborative partnership during the design and development of the above and other projects currently being pursued :

Partner Organizations:

National Library of Medicine: Collaborative Research

Texas Instruments Inc: Collaborative Research

Yale University Yale School of Medicine: Collaborative Research

NASA: Collaborative Research

The Hospital for Special Surgery: Collaborative Research

Conclusions

The NSF-CRCD AWARD # 9980296 entitled "Machine Learning: A Multidisciplinary Computer Engineering Graduate Program " to Texas Tech University has resulted in development of new graduate courses and research program using a top down pyramid learning structure involving faculty advisors, and engineering students at all levels. The research projects were developed in collaboration with federal laboratories, industries, and other academic institutions.

Acknowledgements

The support of the National Science Foundation (NSF-CRCD AWARD # 9980296) and other faculty and student participants are gratefully acknowledged.

Sunanda Mitra

Sunanda Mitra is a Professor of Electrical and Computer Engineering at Texas Tech University. She is also the founder and director of Computer Vision and Image Analysis Laboratory at Texas Tech University. She serves as the principal investigator of the NSF-CRCD program at Texas Tech. Mitra received the Ph. D. from Marburg University, Germany. She is a senior member of the IEEE.