Road Recognition Using Fuzzy Classifiers[1]

J. Shanahan, B. Thomas, M. Mirmehdi, N. Campbell, T. Martin, J. Baldwin
Advanced Computing Research Centre
University of Bristol
Bristol BS8 1UB, UK

Abstract

Current learning approaches to computer vision have mainly focussed on low-level image processing and object recognition, while tending to ignore higher level processing for understanding. We propose an approach to scene analysis that facilitates the transition from recognition to understanding. It begins by segmenting the image into regions using standard approaches, which are then classified using a discovered fuzzy Cartesian granule feature classifier. Understanding is made possible through the transparent and succinct nature of the discovered models. The recognition of roads in images is taken as an illustrative problem. The discovered fuzzy models while providing high levels of accuracy (97%), also provide understanding of the problem domain through the transparency of the learnt models. The learning step in the proposed approach is compared with other techniques such as decision trees, naïve Bayes and neural networks using a variety of performance criteria such as accuracy, understandability and efficiency.

1. Introduction

Fischler and Firschein [1] list learning, and representation and indexing into a large databases of stored knowledge as two of the open issues in computer vision. This situation arose mainly because traditional image understanding focused on techniques from physics, mathematics, psychology, computer science and artificial intelligence that depended tremendously on human input and direction. This usually led to many limitations and endless assumptions on what these techniques could achieve, and usually were labour intensive, limited, and sensitive to change. Examples include knowledge-based approaches such as [2] and [3], and model-based approaches such as [4-6]. More recently people have turned to machine learning as means of building robust, general-purpose systems (for example [7, 8]) with many application areas including image database indexing [7]. Current approaches to computer vision, which use learning, can be differentiated based on the extracted models using the following criteria: effectiveness (accuracy of model on unseen data), understandability (to user or expert in the domain) and evolvability (ability to adapt over time to a changing environment). Most current approaches satisfy understandability or effectiveness, but not simultaneously, while tending to ignore knowledge evolution. For example, Winston's [9] landmark work on symbolic learning in the field of image understanding provided a high level approach using semantic nets to learn object structures from examples and counter-examples (Winston's near-misses) while largely ignoring lower level image processing. Other approaches based upon learning semantic net representations include the classification of hammers and overhead views of commercial aircraft [10]. Michalski et. al. [11] provides some interesting results using a battery of learning approaches: rule-based learn-

---

[1] Communication should be addressed to James.Shanahan@xrce.xerox.com who has recently moved to Xerox Research Centre Europe (XRCE), Grenoble Laboratory, 6 chemin de Maupteruis, 38240 Meylan, France.

ing, neural network learning, and a hybrid of the two. While the application domains of outdoor image classification and the detection of blasting caps in X-ray images of luggage were interesting, there were limited to rather simple uncluttered scenarios. A novel approach to learning the rules of perceptual organisation using fuzzy modelling techniques (resulting in intuitive and transparent models) proposed in [12, 13] may lead to interesting results in the fields segmentation and recognition.

Learning approaches that have provided high levels of accuracy have tended to rely upon extracting models which are dominated by black box modelling and representation techniques: e.g neural network [14-16] or eigenspace-based models [8, 17, 18]. Consequently these approaches provide little or no interpretability or evolvability [19] (not addressed further in this paper). Here we propose a new approach to image understanding, based upon Cartesian granule features [20], that not only provides high levels of accuracy, but also facilitates understanding due to the transparent and succinct nature of the knowledge representation used. The approach is illustrated on a road recognition problem.

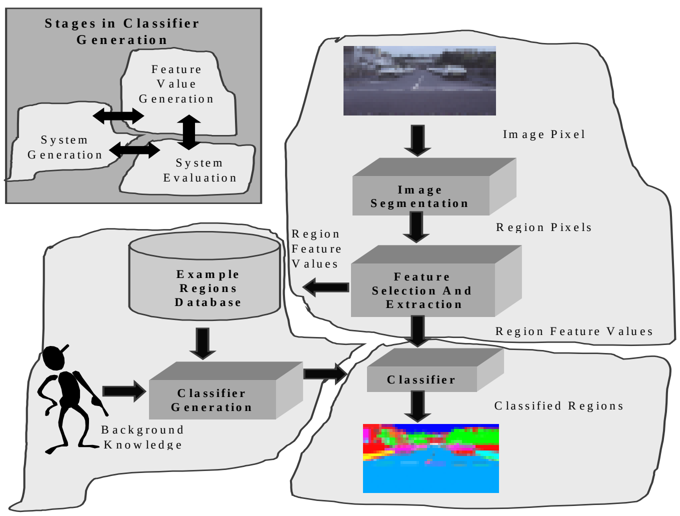


Figure 1: Three stages in classifier generation: Feature Value Generation, System Generation, System Evaluation. Note these stages are iterative.

2. Vision Problem & Proposed Approach

We address the problem of recognising object regions in outdoor scenes. The problem is partitioned into two natural but distinct parts: region segmentation and region classification. Segmentation is achieved using standard image processing approaches (and is not the focus of attention here), and region classification is carried out by a classifier. Each segmented region is described using a variety of features such as colour, location, texture, shape etc.. The main goal of this research is to construct a classifier, automatically from examples, that satisfies the performance accuracy and understandability criteria simultaneously as mentioned in Section 1. In order to meet both criteria we propose to represent the road classifier as an additive fuzzy Cartesian granule feature model that consists of if-then-rules with weighted antecedents whose values are fuzzy sets defined over Cartesian granule features (see for example Figure 3) [20]. Cartesian granule feature based classifiers can be constructed automatically from example data (region feature values) using the G_DACG (Genetic Discovery of Additive Cartesian Granule feature models) constructive induction algorithm [21, 22]. Section 3 gives a brief overview of Cartesian granules features and the G_DACG induction algorithm. Figure 1 presents a block diagram of the proposed approach in terms of three main tasks: feature value generation, system generation, and system evaluation.

3. Modelling with Cartesian Granule Features

Cartesian granule features [20] are a new type of multidimensional feature defined over the Cartesian product of words drawn from the linguistic partitions of the constituent feature universes. They can overcome decomposition error, and also provide model transparency that would facilitate user understanding. Cartesian granules (characterised by fuzzy sets) provide an abstraction of the multidimensional universe by carving it into regions that are drawn together as result of indistinguishability, similarity, proximity or functionality. A fuzzy set can be defined over a Cartesian granule universe as a discrete fuzzy set where each Cartesian granule is associated with a membership value, which is calculated by combining the membership values (using a T-norm such as product or maximum), individual feature values have in the fuzzy sets which characterise the granules.

Figure 2 presents the Cartesian granule fuzzy set induction algorithm. It is presented using an illustrative example of how to extract a Cartesian granule fuzzy set corresponding to car positions in images from example car positions where the top left table corresponds to examples of car positions, corresponding linguistic (fuzzy set) descriptions and least prejudiced distributions (LPDs or probability distributions). Mass assignment theory provides a formal mapping between linguistic descriptions (fuzzy sets) and probability distributions [23]. The top middle graph corresponds to the initial Cartesian granule frequency distribution, where the granule characterisations (i.e. the fuzzy sets) are also shown. The top right graph depicts the Cartesian granule frequency distribution after updating with the LPD corresponding to the value of *40*. The right middle graph shows the Cartesian granule frequency distribution after updating with the LPD corresponding to the value of *60*. The right bottom graph displays the Cartesian granule frequency distribution after counting all the LPDs corresponding to the example car positions. Finally the left bottom graph depicts the corresponding Cartesian granule fuzzy set for car positions in images i.e. a linguistic summary of car positions in images in terms of the words *left*, *middle* and *right.* In this case the Cartesian granule feature is one dimensional in nature for presentation purposes but could be multidimensional. See [21, 24] for full details of the Cartesian granule feature induction process.

In [21] we have shown that systems can be quite naturally described in terms of Cartesian granule features incorporated into rule-based models. Here a region-based fuzzy classifier is constructed automatically from example data (region feature values) using the G_DACG constructive induction algorithm [21, 22]. The G_DACG algorithm discovers good Cartesian granule features (i.e. the feature subsets and the feature universe abstractions). G_DACG is a population-based search algorithm (based on genetic programming), where each node in the search space is a Cartesian granule feature. G_DACG iteratively hones in on good Cartesian granule features based on the evolutionary operations of crossover, mutation and reproduction. Good Cartesian granule features are subsequently incorporated into rule based models (see for example Figure 3). Inference is carried out using evidential reasoning and semantic unification/match of class fuzzy sets and data fuzzy set [23]. Classification corresponds to taking the class associated with the maximum of the inferred results.

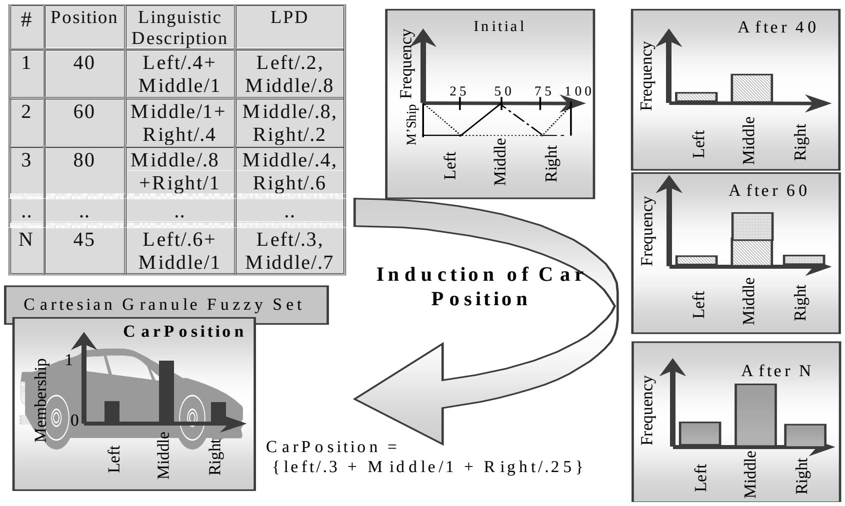


Figure 2: Induction of the Cartesian granule fuzzy set for car positions in images.

4. Feature Selection and Generation

The Bristol Image Database [7, 25] consists of over 350 colour images of a wide range of urban and rural scenes. Figure 5 (lower left quadrant) depicts a typical urban scene in this database. Eighty images of typical outdoor rural scenes were selected from the Bristol image database. Subsequently these images (characterised by intensity) were segmented into *road* and *non-road* regions using the *k*-means segmentation algorithm, where *k* was set to 4. Previous results have shown the *k*-means algorithm to be effective [7] (see also the upper left quadrant of Figure 5). This resulted in 13,628 regions being generated. Feature values were subsequently generated for each region feature. Non-overlapping training, validation and test sets of regions were subsequently generated in a class-wise manner as follows: 70% of data allocated to training, 15% to validation and 15% to testing. Table 1 gives a sample-count breakdown for each class. For the road classification problem each segmented image region was described using a set of over sixty features, comprising of colour, location, orientation, size, shape and texture features. In order to reduce the complexity of the learning process a neural network-based “filter” feature selection algorithm was applied to this feature set [19]. This resulted in ten features been selected as representative features for task of road classification. The first three features correspond to the average luminance and colour differences in a region. The location of the region is expressed as the *X* and *Y* co-ordinates of the region centroid. Orientation is expressed as the sine and cosine of the angle of the principal axis. The next feature corresponds to the principle mode of the PCA (principle component analysis [26]) transformed region boundary description. The last two features arise from the use of a psychophysically plausible model of texture, based upon Gabor filters. In this case the features correspond to two high frequency (128 and 256) isotropic Gabor filters. A full description of all features is presented in [19].

5. Classifier Generation & Application

We applied the G_DACG constructive induction algorithm to the road classification problem. The reduced feature set of ten base features and Cartesian granule features of dimensionality up to three with granularity ranges of (2,..,12) were considered (while parsimony was promoted) thus yielding a search space of over 500,000 nodes. The G_DACG algorithm iterated for fifty generations and at the end of each generation five of the best Cartesian granule features were selected from the current population. The discovered features were then used to form additive Cartesian granule feature rule-based models. Backward elimination was also employed, eliminating extraneous lowly contributing features. The models were evaluated using the test dataset. Table 2 tabulates the results of some of the more interesting ACGF models that were discovered using G_DACG. The models presented in Table 2 were constructed using equal numbers of examples of *Road* and *Not-Road* for training. By equalising the example count across classes a marginal improvement (of less than 1%) in test case accuracy was achieved over learning from the original skewed training set. The results correspond to additive models where the weights have been estimated using semantic discrimination analysis [23]. For example when an additive model consisting of three one-dimensional Cartesian granule features, was formed respectively over the features *Luminance*, *Y-B* (Colour difference) and *Y-Position*, a classification accuracy of 95.5% (after tuning the weights) on unseen image regions was achieved. The feature universes in this model were linguistically partitioned using five words, which are characterised by uniformly placed trapezoidal fuzzy sets with 50% overlap. The additive rule base is presented in Figure 3. The linguistic descriptions, characterised by a Cartesian granule fuzzy sets, corresponding to the *luminance* for *Road* and *Not-Road* classes is presented in close-up detail in Figure 4.

Table 1: Object classifications for each region and corresponding sample counts.

Class No.	Class	# Train examples	# Validation examples	# Test examples
1	Not-Road	8381	1796	1797
2	Road	1157	248	249
TOTAL	13628	9538	2044	2046

Table 2: ACGF models discovered using the G_DACG algorithm.

Dim.	Train % Accuracy	Valid % Accuracy	Test % Accuracy	Optimised Weights	Cartesian Granule Features
1D	92	95	95.5	No	((0 5)) ((2 5)) ((4 5))
2D	94	93.3	96.6	No	((0 5)(2 5)) ((2 5) (4 5))
1D	92	95	95.5	Yes	((0 5)) ((2 5)) ((4 5))
2D	93.9	93.5	96.7	Yes	((0 5)(2 5)) ((2 5) (4 5))

Notice in the additive rule model in Figure 4 that the *Luminance* feature receives a lower weight than the other features involved in the decision-making process. This is mainly because the *Luminance* linguistic summaries do not provide as a good a separation of concepts as the *Y-B* feature for example. Figure 5 presents a Java applet screendump illustrating the results of applying this ACGF model to a k-means segmented image. The results are qualitatively very good from a classification perspective, however the low-level k-means and region growing process has under segmented parts of the image, thus leading to some areas of the image being misclassified.

An additive Cartesian granule model composed of two two-dimensional features give a marginal improvement over the one-dimensional model (see Table 2 for details). The test confusion matrix for this model is presented in Table 3.

6. Comparison and Discussion

The same data and reduced base feature set were used to compare ACGF modelling with standard induction learning techniques. Table 4 summarises the results. The various approaches were evaluated using criteria such as model transparency, performance accuracy, and efficiency of the learning algorithm.

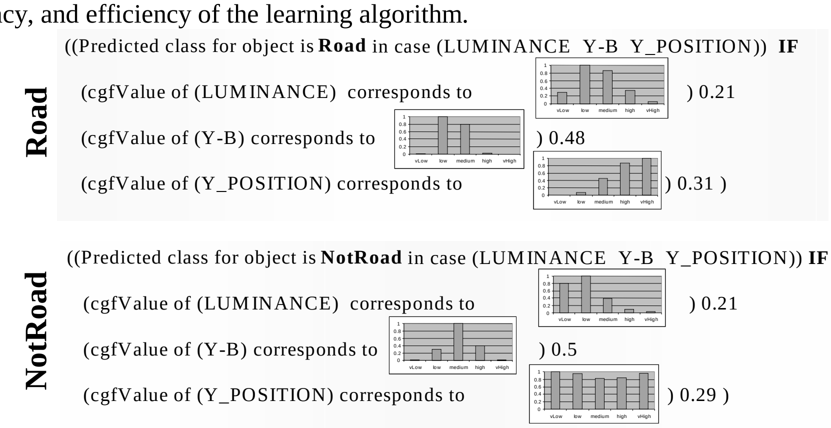
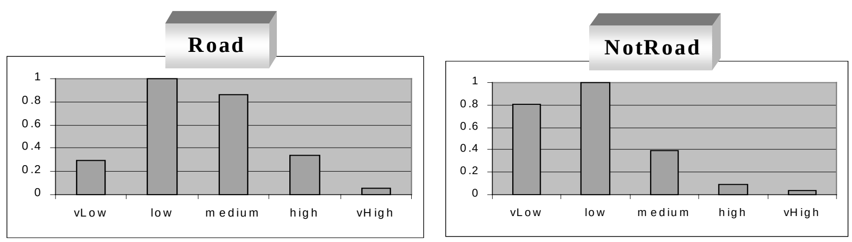


Figure 3: Additive Cartesian granule feature model for road classification.



{vLow:0.29, low:1, medium:0.86, high:0.33, vHigh:0.056}

{vLow:0.8, low:1, medium:0.39, high:0.09, vHigh:0.036}

Figure 4: Linguistic summary, in the form of Cartesian granule fuzzy sets, of *luminance* for Road and Not-Road Classes.

Table 3: Confusion Matrix generated by the discovered 2D model (optimised).

Actual\ Predicted	NotRoad	Road	Total	Class %Accuracy
NotRoad	1767	30	1797	98.3
Road	39	210	249	84.3

6.1 Understandability and Glassbox-ness

One of the primary concerns of intelligent systems is that they should be able to interact naturally with their environment. One of the integral parts of many domains is the human, and consequently the intelligent system (agent) needs to interact with the human. This can be achieved by a variety of means and at many different levels such as a graphic display of trend data. However, one of the most natural forms of communication (and sometimes most effective) is through words. The proposed approach has generated a road classification system that enlightens the user about what a road is, in terms of *luminance* and other feature value descriptions. These descriptions are in terms of words such as *low* and *very low* – generic words in this case, but these could be assigned from a user-defined dictionary and supplemented with hedges such as *very*, *not so much*, etc. and with connectives such as conjunctions and disjunctions. Furthermore the weights associated with each feature inform the user of how important a particular feature is in the inference process.

The induced Cartesian granule feature model facilitates a transition from a low-level object recognition task to a high level understanding task which should greatly simply human computer interaction. This simplification comes from the expression of the knowledge in a form that is almost directly interpretable by the human user. The proposed approach, while facilitating machine learning, may also facilitate human learning and understanding through the generated anthropomorphic models.

With regard to the other approaches examined here, such as the ID3 and C4.5 algorithms, the induced models while being readable tend to be large and consequently makes understanding very difficult. In the case of neural networks and oblique decision trees, the induced knowledge is encoded in vectors of weights (and biases) which may prove difficult for a user to interpret and understand.

A further consequence of readability and understandability is that it will generally increase user's confidence in the system and it can also enhance reliability. For example, the user may enhance the systems reliability by identifying a data deficiency or a variable deficiency.

Figure 5: (top left) The original image, (top right quadrant) k-means segmented image and (bottom left quadrant) the results of region classification using in this case a rule-based ACGF model The regions classified as road are highlighted in grey and the non-road regions are display in black.

6.2 Computational Efficiency

From a classification task perspective, the induced models have similar computational requirements. Learning can be split into two subtasks: structure identification and parameter identification. From a structure identification task perspective (i.e. feature selection, determining a neural network topology, discovering Cartesian granule features) has varying computational requirements for the various approaches considered. These computational requirements are commensurate with the effectiveness of the search techniques used to determine the structure of the induced model. For example, the decision tree approaches (ID3 and OC1) have low computational requirements which arise from the local hill-climbing search technique used. This search technique while facilitating efficient structure identification is vulnerable to local minima. Furthermore, decision tree approaches such as ID3, in order to provide better generalisation, require pruning [30], which can prove to be very expensive in the case of bushy decision trees. On the other hand the G_DACG constructive induction algorithm is computationally intensive which is due to the global, population based search approach used but avoids local minima. The determination of neural network topologies is also computationally intensive.

From a parameter identification task perspective, again the computational requirements vary with the approach used. In the case of ID3 no parameter identification is required. Naïve Bayes parameter identification step has low computational requirements, since the data examples need to be processed only once in order to estimate the class densities. Parameter identification for OC1 and neural networks requires the identification of weights and involves a search through the possible weight space using various search algorithms that offer efficiency commensurate with the effectiveness with the determined solution. Neural networks in the case of this problem are multi-layered and therefore require added computational power than their single-layered oblique decision trees. Parameter identification for additive Cartesian granule feature requires determining the class Cartesian granule fuzzy sets (involves a single pass of the data), and also setting up the class aggregation rules. Additive models can be viewed as a single-layered network where identification of the class aggregation rules reduces to the identification of weights and also of the filter function (not presented here due to space restricetions). These identification tasks are performed independently and consequently are significantly more efficient than the parameter identification of a multi-layered neural network.

In the case of additive Cartesian granule feature models, the system identification step is not just concerned with identifying a model that provides high performance accuracy (the goal of most other induction algorithms), but is also concerned with identifying a model is glassbox in nature. This issue of identifying glassbox models, while having extra computational requirements, is compensated by the identification of models that facilitate understandability.

Table 4: Comparison of results obtained using a variety of machine learning techniques on the road classification problem.

Approach.	# Features used	% Accuracy
Additive Cartesian granule feature model	3	96.7
Naïve Bayes [27]	10	96.2
Oblique Decision Trees [28]	10	94.7
Feed forward Neural Net [29]	10	97
ID3 or C4.5 [30]	10	92.75

6.3 Feature value representations

All input features for the road classification problem are continuous in nature. The values of these features are single numbers, which in the case of some features correspond to simple statistical measures such as the average. For example the *luminance* value for a region corresponds to average pixel luminance value across that region. In this case of the road classification problem such features prove adequate in modelling the problem, but in the case of a more difficult multi-class problem more detailed feature values may be necessary. Single numeric values such as the measures used here can be susceptible to noise, and generally lead to high data requirements for learning. An alternative, and possibly promising, approach is to linguistically summarise the pixel values using a one-dimensional Cartesian granule feature i.e. generate a linguistic histogram. Other features that may benefit from linguistic summaries include the texture features, the colour difference features, location feature etc. Linguistic summaries of feature values provide more information to discriminate amongst different classes while also combating the curse of dimensionality.

6.4 Alternative features

Current work describes concepts in terms of their own attributes, whereas more succinct and possibly easier to understand concept definitions can be acquired where objects are described in terms of other objects. For example an object could be defined as *similar* to another object. In the case of induced Cartesian granule feature models, class rules can be embellished by hand by adding in further conditions that a class should satisfy. For example, a reasonable condition for the road class is that "*cars should be above road"*, where the user provides the condition and also a definition for *above* (or alternatively *above* could be extracted from example data by taking the difference in *y-position*.

7. Conlusions

A new approach to object recognition, based upon a Cartesian granule feature classifier, has been proposed that facilitates the transition from recognition to understanding. The approach was illustrated on a road classification problem, yielding high levels of accuracy (97%) and very understandable models. The approach, when compared with decision tree approaches, naïve Bayes and neural networks provided simpler models with better accuracy that takes a little longer to discover. The extra discovery time needed is mainly due to the search for a transparent model. Potential applications of the proposed approach include autonomous vehicle navigation systems and landmine detection, where magnetic resonance images of test sites could be used to train an additive Cartesian granule feature model which would subsequently detect landmines in a consistent and effective manner reducing the number of false alarms that can prove expensive. Other envisioned applications include content based image retrieval systems (CBIR). CBIR is an area which relies heavily on human-computer interaction, where interaction requires understanding, and thus would greatly benefit from the glassbox approach proposed here.

Acknowledgements

This work was supported by DERA Grant 92W69.

Bibliography

1. Fischler, M.A. and O. Firschein, eds. *Readings in computer vision: issues, problems, principles, and paradigms*. . 1987, Kaufmann Publishers, Inc.: Los Altos, CA, USA. 765-768.
2. Draper, B.A., R.T. Collins, J. Brolio, A.R. Hanson, and E.M. Riseman, *The Schema system.* The International Journal of Computer Vision, 1989. **2**: p. 209-250.
3. Strat, T.M., *Natural object recognition*. 1992, New York, USA: Springer-Verlag.
4. Mirmehdi, M., P.L. Palmer, J. Kittler, and H. Dabis, *Feedback control strategies for object recognition.* IEEE Transactions on Image Processing, 1999. **(To appear)**: p. 19.
5. Grimson, W.E.L. and T. Lozano-Perez, *Model-based recognition and localization from sparse range or tactile data.* Int'l Jrnl of Robotics Research, 1984. **3(3)**: p. 3-35.

6. Brooks, R.A., *Model-based three-dimensional interpretations of two-dimensional images*, in *Readings in computer vision: issues, problems, principles, and paradigms*, M.A. Fischler and O. Firschein, Editors. 1987, Kaufmann Publishers, Inc.: Los Altos, CA, USA. p. 360-370.
7. Campbell, N.W., B.T. Thomas, and T. Troscianko, *Automatic segmentation and classification of outdoor images using neural networks.* International Journal of Neural Systems, 1997. **8**(1): p. 137-144.
8. Murase, H. and S.K. Nayar. *Learning and recognition of 3D objects from appearance*. in *IEEE 2nd Qualitative Vision Workshop*. 1993. New York, NY.
9. Winston, P.H., ed. *The Psychology of Computer Vision*. . 1975, McGraw-Hill: USA.
10. Connell, J.H. and M. Brady, *Generating and generalising models of visual objects.* Artificial Intelligence, 1987. **34**: p. 159-183.
11. Michalski, R.S., A. Rosenfeld, Z. Duric, M. Maloof, and Q. Zhang, *Learning patterns in images*, in *Machine Learning and Data Mining*, R.S. Michalski, I. Bratko, and M. Kubat, Editors. 1998, Wiley: New York. p. 241-268.
12. Ralescu, A.L. and J.G. Shanahan. *Line structure inference in fuzzy perceptual grouping*. in *NSF Workshop on Computer Vision*. 1995. Islamabad, Pakistan: Springer.
13. Ralescu, A.L. and J.G. Shanahan, *Fuzzy perceptual organisation of image structures.* Pattern Recognition, 1999. **To Appear**: p. 20.
14. Wood, M.E.J., N.W. Campbell, and B.T. Thomas. *Searching large image databases using radial basis function neural networks*. in *International Conference on Image Processing and its Applications*. 1997. London, U.K.
15. Campbell, N.W., W.P.J. Mackeown, B.T. Thomas, and T. Troscianko, *Interpreting Image Databases by Region Classification.* Pattern Recognition, 1997. **30**(4): p. 555-563.
16. Mukunoki, M., M. Minoh, and K. Ikeda. *Retrieval of images using pixel based object models*. in *IPMU*. 1994. Paris, France.
17. Kosako, A., A.L. Ralescu, and J.G. Shanahan. *Fuzzy techniques in Image Understanding*. in *ISCIE Joint Conference for Automatic Control, Osaka, Japan*. 1994.
18. Turk, M.A. and A.P. Pentland. *Face recognition using eigenfaces.* in *IEEE Conf. on Computer Vision and Pattern Recognition*. 1991.
19. Shanahan, J.G., M. Mirmehdi, J.F. Baldwin, B.T. Thomas, N.W. Campbell, and T.P. Martin, *Transitioning from recognition to understanding in vision using additive Cartesian granule feature discovery models.* Journal paper in preparation, 1999: p. N/A.
20. Baldwin, J.F., T.P. Martin, and J.G. Shanahan. *Modelling with words using Cartesian granule features*. in *FUZZ-IEEE*. 1997. Barcelona, Spain: University of Bristol, UK.
21. Shanahan, J.G., *Cartesian Granule Features: Knowledge Discovery of Additive Models for Classification and Prediction*, . 1998, PhD Thesis, Dept. of Engineering Maths, University of Bristol, Bristol, UK.
22. Baldwin, J.F., T.P. Martin, and J.G. Shanahan, *System Identification of Fuzzy Cartesian Granule Feature Models using Genetic Programming*, *Lecture notes in Artificial Intelligence (LNAI 1566)*, A.L. Ralescu and J.G. Shanahan, Editors. 1998, Springer. p. 26.
23. Baldwin, J.F., T.P. Martin, and B.W. Pilsworth, *FRIL - Fuzzy and Evidential Reasoning in A.I.* 1995: Research Studies Press(Wiley Inc.), ISBN 086380159 5.
24. Baldwin, J.F., T.P. Martin, and J.G. Shanahan, *Knowledge discovery with words using Cartesian granule features.* (To Appear) Jrnl of Computational Intelligence, 1999. **N/A**: p. 40.
25. Mackeown, W.P.J., P. Greenway, B.T. Thomas, and W.A. Wright. *Contextual Image Labelling with a Neural Network*. in *IEE Vision, Speech and Signal Processing*. 1994.
26. Jolliffe, I.T., *Principal Component Analysis*. 1986, New York: Springer.
27. Good, I.J., *The estimation of probabilities: an essay on modern Bayesian methods*. 1965: M. I. T. Press.
28. Murphy, S.K., S. Kasif, and S. Salzburg, *A system for induction of oblique decision trees.* Journal of Artificial Intelligence Research, 1994. **2**: p. 1-33.
29. Rumelhart, D.E., G.E. Hinton, and R.J. Williams, *Learning internal representations by error propagation*, in *Parallel Distributed Processing (Volume 1)*, D.E. Rumelhart and J.L. McClelland, Editors. 1986, MIT Press: Cambridge, USA.
30. Quinlan, J.R., *C4.5: Programs for Machine Learning*. 1993, San Mateo: Morgan Kaufmann.