Using genetic algorithm feature selection in neural classification systems for image pattern recognition

Implementación de selección de características con algoritmos genéticos en clasificadores neuronales para reconocimiento de patrones en imágenes

M. R. Gamarra A.[1] and C. G. Quintero M.[2]

ABSTRACT

Pattern recognition performance depends on variations during extraction, selection and classification stages. This paper presents an approach to feature selection by using genetic algorithms with regard to digital image recognition and quality control. Error rate and kappa coefficient were used for evaluating the genetic algorithm approach Neural networks were used for classification, involving the features selected by the genetic algorithms. The neural network approach was compared to a K-nearest neighbour classifier. The proposed approach performed better than the other methods.

Keywords: Feature vector, genetic algorithm, neural network, pattern recognition.

RESUMEN

El desempeño en el reconocimiento de patrones depende de las variaciones en las etapas de extracción, selección y clasificación. En el siguiente artículo se presenta un enfoque para selección de características utilizando un algoritmo genético aplicado a procesos de reconocimiento y control de calidad en imágenes. Para evaluar la propuesta con algoritmos genéticos se utilizan dos funciones de evaluación: tasa de error y coeficiente Kappa. Se implementan redes neuronales en la clasificación usando las características seleccionadas por el algoritmo genético. La red neuronal se compara con el clasificador de los k-vecinos. Los resultados obtenidos muestran un mejor desempeño del sistema propuesto frente a otros métodos.

Palabras clave: Algoritmos Genéticos, Vector De Características, Redes Neuronales, Reconocimiento de Patrones.

Introduction

Technological advances in our society have enhanced industrial production automation and the need for handling information in fields of great importance. This trend has earned pattern recognition a central place in research and engineering applications. Pattern recognition has become an integral part of most machine vision systems' output regarding decision-making (Theodoridis S. & Koutroumbas, K, 2009) for identifying a given object and performing operations on it.

Pattern recognition has been developed using classical statistical methods such as principal component analysis (PCA) (Song F,

[1] Margarita R. Gamarra A., M.Sc. Ingeniería Electrónica Universidad del Norte. Ingeniería Electrónica – Ingeniería Eléctrica Universidad del Norte. Affiliation: Lecturer in the Electrical and Electronics Engineering Department, Universidad del Norte. Electronics and Telecommunications Research Group Coordinator, Universidad Autónoma del Caribe, Barranquilla (Colombia).
E-mail: gamarram@uninorte.edu.co

[2] Christian G. Quintero M., Ph.D. Tecnologías de la Información, Universidad de Girona. Affiliation: Full time professor in the Electrical and Electronics Engineering Department and Program Coordinator for the Master's degree in Electronics Engineering at Universidad del Norte, Barranquilla (Colombia).
E-mail: christianq@uninorte.edu.co

Guo Z, 2010) and independent component analysis (ICA) (Ekenel H, Sankur B, 2004; Liu J, Wang G, 2010). However, the three processes involved in pattern recognition can be developed by using computational intelligence, thereby providing an autonomous system having learning ability and suitable classification results, according to recent research (Mitra S. & Sankar K, 2005; D. G. Stavrakoudis, 2010; Vilches E, 2006; Pazoki Z, Farokhi F, 2010). This work was focused on the selection stage by using a genetic algorithm (GA) approach; the classification stage involved a neural network-based approach. The proposed GA approach's performance was evaluated by using fitness functions, classifiers and compared to other research.

A GA optimises a fitness function and (in this case) determines the best subset of features. However, in other work (i.e. Daza G.S, Sánchez L.G, 2007; Rivera J.H, Castellanos C, 2007; Changjing S, Barnes D, 2009) selection has been limited to a number of specific characteristics, one-dimensional signals or face recognition tasks whilst GA has been used to tune other techniques' parameters.

Some approaches have proposed effective feature selection, aimed at reducing feature dimension and having minimum classification errors. GA have been applied to digital images regarding recognition and quality control, using different features (texture, statistics, colour and other descriptors and transforms). Two

experiments have been proposed for evaluating a GA-based approach, using error rate (Fuentes G, et. al., 2006) and Kappa coefficient (Vieira S, et. al., 2010), according to the fitness function.

An artificial neural network (ANN) can be used for classification by using the features selected for the GA. A neural network (NN) approach is appropriate for classification because input-output relationships in complex and non-linear conditions can be established on it.

Feature selection methods aim to extract features which are important in classification. Orozco (2005) used a wavelet feature selection-based method involving GA which addressed the classification system's three stages: feature extraction, feature selection and classification. PCA was also used as the standard technique for recognition and classification experiments. The experiments showed that the nearest feature line (NFL) was most appropriate, considering its accuracy, the amount of information (variance) needed and the number of distance calculations.

Fuentes (2008) proposed a face recognition method divided into two stages: feature extraction by discrete wavelet transform (DWT) (reducing image dimensionality) and classification of feature vectors through a multilayer perceptron NN. The latter type of network was able to solve highly nonlinear problems and enabled non-parametric classification of facial features. The feature vectors extracted by DWT were suitable for face recognition. Daubechies wavelet filters of order 3 and order 12 (Symlet) had the highest recognition rates. A combination of NN and DWT provided a robust and efficient model.

Recent research has applied hybrid algorithms combining several intelligent techniques to improve pattern recognition. Umamaheswari (2006) proposed using GA and NN to detect human faces and locate eyes in real time. The hybrid algorithm focusing on face recognition was more efficient than current digital image processing techniques, such as wavelets.

Jeong et al., (2010) presented a feature selection method using a backward selection algorithm reducing the total set to the most relevant for classification by using a Kappa coefficient. The method was for classifying defects in steel from images to improve quality inspection system performance.

The above results have shown that using intelligent algorithms for pattern recognition has led to better classification performance. However, the number of classes has been small (2 or 3), the features have been specific, applied to one-dimensional signals and 2D signals have been limited to face recognition.

The proposed research was thus aimed at implementing a method for effective feature selection based on GA and classification with ANN. Two fitness functions were used and compared (error rate and Kappa coefficient). NN was compared to the K-nearest neighbour (KNN) classification algorithm during the classification stage. The proposed approach was compared to other face recognition methods (Fuentes Gibran, 2008). Images of car parts (classification and quality control), fruit and faces were used for pattern recognition, taking objects' characteristics and different classes into account, reducing feature size and minimising the error rate.

The proposed approach

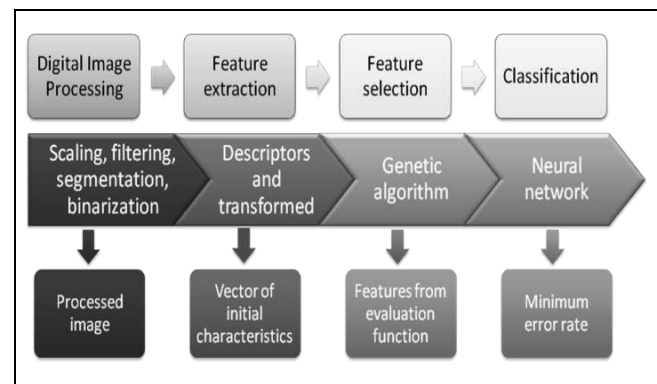Figure 1 outlines the pattern recognition process.

Figure 1. A diagram of the proposed approach

The first stage referred to techniques used for image processing (filtering, resizing and binarisation).

The second stage involved extracting a set of characteristics (shape descriptors, statistical descriptors and coefficients) forming the basis for pattern recognition. If the characteristics of the objects to be classified varied within a class, the feature extraction stage required a robust image. Wavelet transform coefficients were extracted at this stage for face images; different descriptors were used for other cases.

The initial feature vector was reduced to the most effective one in the pattern to be classified during the next stage. The aim of feature selection is to use fewer features to achieve the same or better performance (Amine A et al, 2009).

A GA approach was used in the feature selection stage, assuming that pattern recognition is complex and involves several stages (Wang Y, Fan K, 1996). Most optimisation theory algorithms usually work well if the objective function has all the prerequisites. However, GA is a search algorithm based on population and is the best method for linear objective functions in a search space limited by linear constraints (Wang Y, Fan K, 1996).

The fitness function is the most important part of a search algorithm. The purpose of search strategy is to find a subset of features to optimise such function (Amine A et al, 2009). The percentage of misclassification was thus used as a fitness function. The following expression was considered for GA minimisation of classification error and reduce the size of features for optimising the value:

$$V \min = \%error + \frac{Feature\_subset}{k * Total\_features} \qquad (1)$$

where *Feature_subset* expressed the reduced feature dimension, *Total_feature* was the original dimension and *k* was placed to give priority to reducing the percentage of error. This factor minimised the term involving the amount of features; it was experimentally determined and selected 10 for metal parts and fruit and 1 for face recognition.

The GA applied in this approach generated an initial binary vector where each bit was associated with a feature: if bit value was 1, the property was taken into account, and if bit value was 0 it was not taken into account. This feature vector was applied to the training and test images of the KNN classifier. Input for this classifier consisted of training image values, having a correspondence to the class to which they belonged. Output consisted of test image classes. The classification error was calculated as the percentage of wrong classes. This was repeated until the classification error was found to be minimal.

The Kappa coefficient was used for the fitness function to consolidate the research results; this coefficient has been used in several investigations (Alejo *et al.*, 2009; Landis, J. & Koch, G, 1977; Jeong *et al.*, 2010; Vieira *et al.*, 2010). This value indicated the percentage of agreement regarding classification when the random part had been eliminated, i.e. it showed how much the classification system had improved regarding a random classification (Alejo *et al.*, 2009). The error rate was changed by the Kappa definition for when designing the second experiment (J.A. Cohen, 1960):

$$K = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)} \qquad (2)$$

where $P_r(a)$ was the percentage of hits and $P_r(e)$ the coefficient due to chance. If classes were classified in complete agreement, then $K = 1$. If the scores were lower than those given by chance $K < 0$. The best performing classifier should then have had a larger K (Vieira *et al.*, 2010).

The GA approach reduces the number of features in addition to increasing K, just as in the error rate case. The second term of equation (1) was included in the value to be optimised. The KNN classifier was used in this experiment again.

The last stage consisted of classification. The function of a classifier is to divide the space of characteristics representative of each pattern associated with a single class or label in many regions (Gómez, L, 2010).

An ANN was used in this stage; one of the main characteristics of an ANN is that they can learn complex nonlinear input-output relationships, adapting to the data (Basu J, Bhattacharyya D, 2010). An ANN can modify the parameters, depending on performance; this makes them suitable for classification (Basu J, Bhattacharyya D, 2010).

The ANN proposed here took the feature vector selected by the GA as input and, by training images, it provided interconnections so that it could perform classification functions with minimal error.

Experimental framework

Databases and test parameters

Two experiments using MATLAB were designed to evaluate the proposed approach. The first used the percentage of misclassification in the fitness function and the second searched for the function that aimed to maximise the Kappa coefficient. The three databases (car parts, fruit and faces) were used in both experiments.

Tests also involved using another optimisation algorithm known as sequential forward selection (SFS) which has been widely used in the literature (Fuentes *et al.*, 2006), for each case being studied. The results were compared to GA and SFS in both experiments. ANN results from the classification stage were also presented and results compared to Fuentes Gibran (2008) whose work presented several feature selection methods.

Twenty neurons were used in the hidden layer, except for housing quality control which involved 30 and face recognition stage involving 40. This characterisation was carried out experimentally.

The repeated holdout validation method was used for each case study, allowing classification error to be estimated. This method consisted of dividing the available data in the training and validation sets by 30% of the data for testing and 70% for training, as suggested in Fuentes (2006).

Twenty repetitions were made for each case when designing experiments with the training and test sets for each database. Training and test data sets were updated on each repetition, keeping the ratio 70%-30% of the total set of images. Experiments involved using a 90% confidence interval.

Case 1: Car parts

Two databases consisting of images of metal parts in JPEG format had been previously collected by Quintero (2010). One contained three types of parts (gear wheel, housing and central gearbox), aimed at identifying the three classes. The second database contained a single type of part but in different states, aimed at classifying each part into one of four possible states (rejected, accepted, defective or reprocessing part).

The original feature set was different for each database, because they dealt with different processes: recognition and quality control; 13 features were taken in the first case and 17 in the second one, including descriptors of form, statistical moments and additional information regarding the parts.

Figure 2. Examples of a dataset: a gear wheel, housing and gearbox

To train and test the classifier, 32 images of each class available in the database were used. Only the main view of each part was used for recognition and quality control to facilitate the above.

Given the number of parts for training, the classifier parameters were set to 10 nearest neighbours for cataloguing a part for a class. This involved around half the training samples. The GA parameters were determined experimentally (Grefenstette, J, 1986) by selecting the following: population size 10, number of generations limit 100, elite population 2 and crossing factor 0.2.

Case 2: Fruit

Fruit images were taken from the Amsterdam Library of Object Images (ALOI) database for the second case study; such images were in PNG format. Each fruit had several shots by changing the rotation every 5°; there were also two different types of images for each fruit (Figure 3). The aim was to recognise 4 different classes (apple, lemon, orange and strawberry). The uniform background and correct lighting facilitated image processing
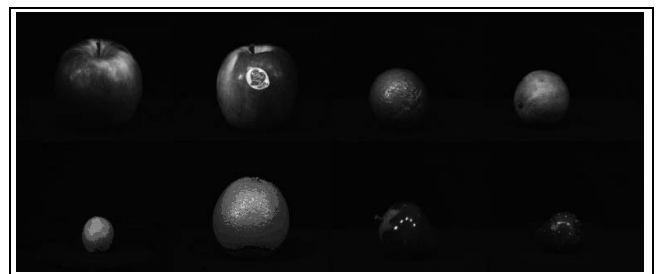

Figure 3. Examples of a dataset for fruit recognition

The original feature set had 23 values, corresponding to greyscale shape, texture and statistical moment descriptors. Statistics were also taken regarding each RGB plane.

144 images of each class (available in the database) were used to train and test the classifier. Given the number of training images, the classifier parameters set 70 nearest neighbours to classify the fruit into a class. The GA parameters were selected in the same way as for the first database.

Case 3: Faces

Images were taken from the AT&T Laboratories (Cambridge) ORL database for testing face recognition; this presents variations in facial gesture and position. Figure 4 shows examples of face images.

The original feature set involved low frequency band wavelet coefficients, containing information relevant for recognition between classes.

Wavelet transform is commonly used in face recognition, because of the information provided for this type of process. According to the results obtained in Fuentes (2008), the Daubechies of order 3 and Symlets order 12 were selected from different wavelet families. Both had a lower percentage error rate and their properties were very similar.

Figure 4. Examples from the faces dataset

To train and test the classifier, 10 images of each person were used for identification and 20 people for recognition (all available in the database). As the number of samples per class was smaller than the number of classes, 10 pictures of each individual were taken for validation, including images used in training.

As 70% of the images were taken for training, 7 neighbours were used in the classifier. The GA was set for the following parameters: population size 80, number of generations limited to 100, elite population 2 and crossing factor 0.2. Population size was increased because there were more baseline characteristics.

Results

Feature selection

The GA-based approach for each database involved several runs in all datasets.

Car parts dataset

All runs in the car parts database regarding recognition and quality control gave the same value, selecting a set of features for minimum error. An average 3.7% classification error was only obtained by selecting feature number 7 (part perimeter).

Several runs were made with gear wheel quality control; 3 features were selected for a minimum error (mean, distance between holes and number of teeth). These characteristics agreed

with those used for quality control by Quintero (2010). Average classification error was 6.25%.

GA were also used to verify gearbox and housing parts' quality control, classifying classes corresponding to the parts' different states (accepted, rejected, defective and reprocessing).

Average housing classification error was 3.89%, using 92 training images and 27 test images. The three features selected to give a minimum error were: number of holes, diameter and the third statistical moment.

The algorithm determined that the effective features regarding gearbox quality control were the mean and percentage rusted area of the part. A 2.36% classification error was obtained for the test images when using these two features.

An additional test was made for recognising parts; the number of classes was increased to 4 (gear wheel, housing and two different views of the gearbox piece) to determine whether the number of selected features changed (selected features were diameter and perimeter). Average classification error was 3.33% using only these two characteristics.

Different features were selected although the three types of part had similar patterns regarding quality control for each case. The selected features were also different, similarly to including another class in recognition (i.e. descriptors of shape, texture, statistical moments and invariant moments).

Error rates were also analysed for an increased number of classification features and a case of quality control, taking the best results for each subset (Fig. 5.a, Fig. 5.b).
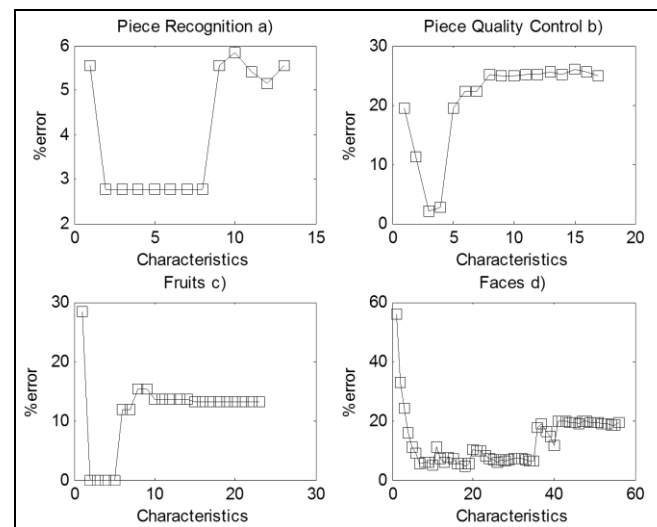

Figure 5. Error rates cf number of characteristics for: a) recognition of parts, b) quality control of parts, c) fruit and d) faces

As expected, increasing the number of features increased the percentage of error regarding recognition since there were similar features for classes, thereby hampering more characteristics. Likewise, using fewer features increased the percentage of error.

Fruit dataset

Only two features were selected from an initial set of 23 (minor

axis length and third invariant moment). Figure 5.c shows the different error rates, when the number of features used in recognition were increased.

Increasing the number of features meant that error rate also increased because some characteristics were similar to several classes, thereby hampering their recognition. An average 1.7% classification error was obtained in tests by using 400 training images and 176 for testing.

Faces dataset

The original set of features was large because it concerned a series of wavelet coefficients. These coefficients led to good recognition of faces. Thus, low error rates were obtained, even by using more features (Fig. 5.d). The average error rate was 7.3% using only 7 coefficients.

Experiment 1

A vector of zeros for features was taken as starting point with the SFS algorithm. Error rates were compared to reduced features and to all of these for each case evaluated and the number of features used.

Figure 6 shows a comparative graph of error rates with reduced features by using a GA, SFS algorithm and all the characteristics for each case evaluated.
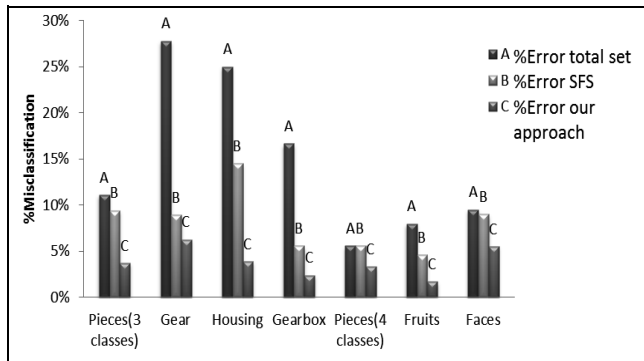

Figure 6. Feature reduction results for different case studies

Figure 7 presents a comparative graph of the number of features selected by the SFS optimisation algorithm and the GA.
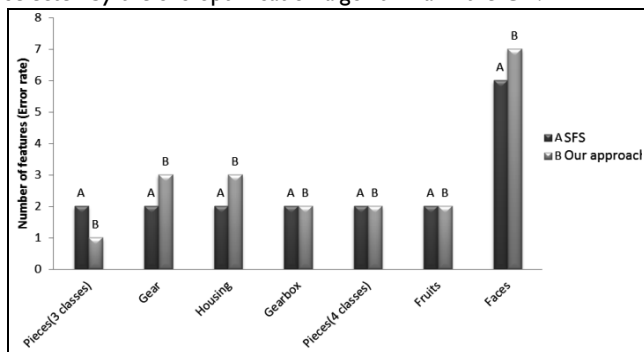

Figure 7. The number of selected features from the SFS and GA

GA error rates were lower than the results for the SFS algorithm, according to Figures 6 and 7; however, the SFS algorithm had a larger reduction in the number of selected features in some cases.

Likewise, tests were performed with an equal amount of features selected by the GA, but with a different set. The results for each case study are shown in Figure 8, with the minimum values obtained for different subsets. These subsets were randomly selected.

GA not only reduced the size of the features selected but also provided a lower percentage of error (Fig. 8).
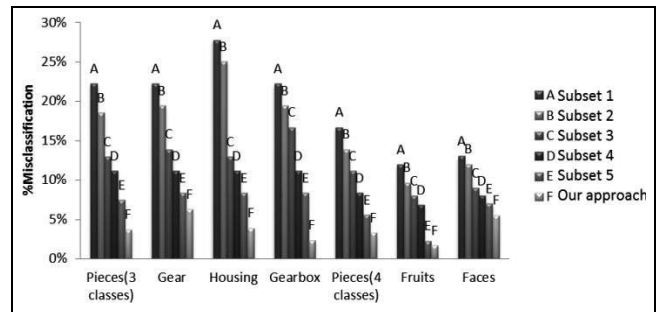

Figure 8. Results with other subsets of features for different cases

The average error rate and the confidence interval for each case study is shown in Table 1, taking into account that 20 repetitions were performed to reduce the interval and set a 90% confidence level. The results were compared with the KNN classifier, used in the GA fitness function and using NN, feeding on the features selected by the GA.

Using ANN in the classification stage improved recognition rates in fruit, faces, housing and gearbox by choosing the features selected by the GA in Table 1. Confidence intervals overlapped regarding gear wheels and the recognition of parts with 3 and 4 classes; it was thus considered that there was no significant difference in the results of two classifiers, for these cases.

Table 1. Confidence interval error rate for several cases

Case study	%Error (K-neighbours)	Confidence interval	%Error (neural networks)	Confidence interval
Parts (3 classes)	3.70%	2.5%-4.8%	2.85%	2.4%-3.2%
Gear wheel	6.25%	5.7%-6.7%	6.23%	5.6%-6.8%
Housing	3.89%	3.4%-4.4%	2.95%	2.6%-3.3%
Gearbox	2.36%	1.6%-3.2%	0.34%	0.1%-0.5%
Parts (4 classes)	3.33%	2.5%-4.1%	2.78%	2.2%-3.3%
Fruit	1.70%	1.1%-2.2%	0.31%	0.1%-0.5%
Faces	7.3%	6.8%-7.7%	4.2%	4%-4.4%

Experiment 2

The GA selected the same subset of features as in the first experiment, except for face recognition which selected different characteristics.

Figure 9 shows that GA had higher Kappa values for the second experiment, compared to the SFS algorithm and with the full set of features.

The results were only the same as the SFS algorithm for gearbox quality control and the recognition of three parts. This was because a fitness function was used (Kappa coefficient) which improved SFS algorithm results, compared to the percentage of error as fitness function.

Landis and Koch (1977) proposed an assessment regarding the degree of agreement of classifiers based on the Kappa value; values greater than 0.6 would have indicated substantial agreement between predicted and observed data. According to the results shown in Figure 11, the GA gave values above 0.9.
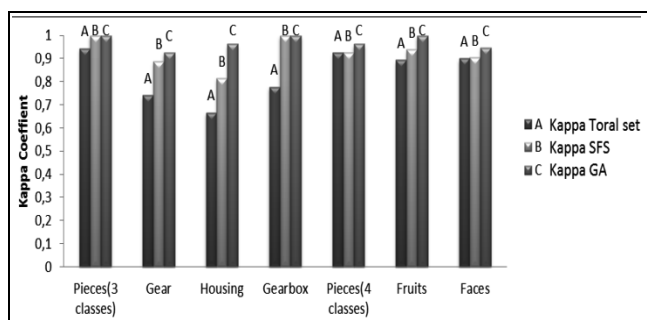
Figure 9. Kappa coefficient experiment results regarding feature reduction for several cases.

Figure 10 shows that the SFS algorithm in the second experiment selected a lower or equal number of features compared to the GA. These results were similar to those obtained in experiment 1 which used fewer features but had higher percentage of error.
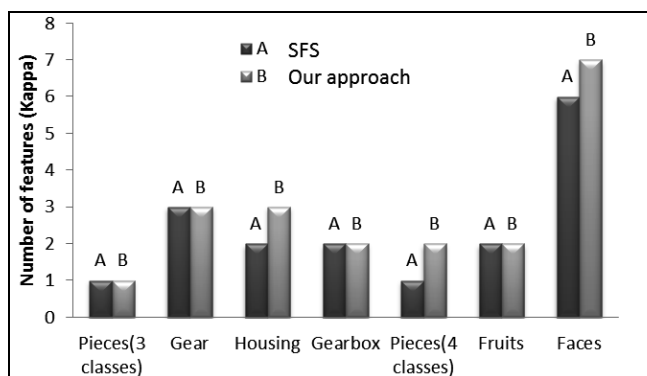

Figure 10. The number of features selected from SFS and GA. Kappa coefficient experiment

By taking the same amount of some subsets of features as those selected by the GA, the method had better results (Figure 11).

The GA selected the same characteristics as in the first experiment; classification results obtained with NN are shown in Table 1.
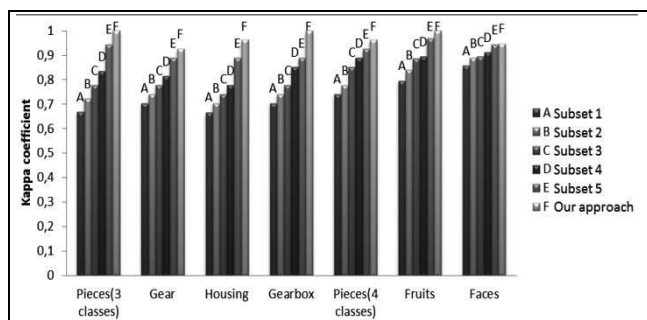

Figure 11. The results regarding other subsets of features for several cases. Kappa coefficient experiment.

Comparison with other approaches

Section 1 mentioned results of research regarding image pattern recognition. Wavelet transform has been used for feature selection and NN classification (Fuentes G, 2008); results were compared to other proposed feature selection methods concerning faces (Aguilar *et al.*, 2007) (Table 2).

The same parameters described in Fuentes (2008) were used for comparison regarding performance on face recognition experiments with the ORL database.

Table 2 shows that the proposed method significantly reduced the number of features used, having a higher recognition rate. Discrete wavelet transform was used in the feature extraction stage (Fuentes *et al.*, 2008) and a NN as classifier, with the same parameters.

However, a GA was used with the error rate as a fitness function to reduce the number of wavelet coefficients to be used by the ANN.

Table 2. Comparison of performance with other selection methods

Method	Number of features	% recognition
Eingenfaces	150	83
DCT	160	90.67
FG	108	85.67
Discrete Walsh transformer	80	75.33
Eingenfaces + FG	258	92
Eingenfaces+ FG + discrete Walsh transformer	338	86.7
DWT (level 3)	168	87.33
DWT (level 4)	42	88.67
Proposed method	11	94

Conclusions

Given these results, it was concluded that the feature selection stage using a GA-based approach eliminated redundant information, thereby reducing the number of features for recognition and leading to lower error rates.

The influence of the classifier on GA performance was noted in tests. The features in each run were quite different for a single neighbour, although sometimes having the same minimum value of error. When the number of neighbours was increased, it was observed that algorithm performance became uniform and reduction more effective.

The graphs showed a significant reduction in the number of features, with misclassification being very low compared to errors by using all the features or more of them.

The proposed algorithm managed to reduce the feature space and ensured optimal selection of features because, in spite of using a minimum number of them, error rates lower than such performed by the GA could not be achieved with other subsets.

The relationship between the number of features and the error rate showed that increasing feature sets meant an increase in classifier percentage of error because the classes tended to have similar values regarding certain properties. Error rates remained low for face recognition for an increase in the number of wavelet coefficients because these coefficients provided good identification of faces.

The SFS algorithm generated a vector having minimum error rate, but did not optimise the fitness function. The genetic algorithm thus generated population crossing instead of growing along a gradient, avoiding concentration in a local minimum and then reaching the best solution.

The Kappa coefficient was also implemented in the GA fitness function for analysing the feature selection results. This measurement had also lower error rates and reduced number of features.

The K-neighbour classifier has been seen to be efficient and easy to implement; it was used for the fitness function in GA selec-

tion, leading to low percentages of error being obtained. However, NN-based approaches were used in the classification stage, thereby improving class recognition. The GA-based approach selected features effectively and NNs reduced the percentage of error.

The classification stage (involving features selected by the GA) improved recognition rates when using a NN approach (Table 1). Only in some case regarding parts recognition and gear wheel quality control were the results similar to those shown by the KNN classifier.

When comparing the proposed methodology's performance to that of others reported in the pertinent literature regarding face recognition, the results were better with a GA-based approach in the selection stage and NN in classification as this greatly reduced the features used, thereby obtaining higher recognition percentage. However, the proposed approach should include an additional step after obtaining the wavelet coefficients involving GA selection to minimise the number of features fed to the NN.

This research involved comparative analysis of classifiers, taking error rate into account. However, further study should analyse computational cost and classifier execution time, taking into account that pattern recognition can be applied in real time and the hardware involved must involve minimal complexity.

Acknowledgment

References

Aguilar G, et. al., Face Recognition Algorithm Using the Discrete Gabor Transform, Proceedings of the 17th International conference on Electronics, Communications and Computers, 2007.

Alejo R, et. al., Redes Neuronales Artificiales y Distribuciones no Balanceadas, Programación matemática y software, Vol. 1, 2009.

Amine A, et. al., GA-SVM and Mutual Information based Frequency Feature Selection for Face Recognition, GSCM-LRIT, Faculty of Sciences, Mohammed V University, B.P. 1014 Rabat, Morocco, 2009.

Amsterdam Library of Object Images (ALOI) http://staff.science.uva.nl/~aloi/

AT&T Laboratories Cambridge. The ORL database of faces. http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html, 1992-1994.

Basu J, Bhattacharyya D, Kim T, Use of Artificial Neural Network in Pattern Recognition, International Journal of Software Engineering and Its Applications Vol. 4, No. 2, 2010.

Changjing S, Barnes D, Giang S, Effective Feature Selection for Mars McMurdo Terrain Image Classification, Ninth International Conference on Intelligent Systems Design and Applications, Department of Computer Science, Aberystwyth University, UK, 2009.

D. G. Stavrakoudis, J. B. Theocharis, Employing Effective Feature Selection in Genetic Fuzzy Rule-Based Classification Systems, in: Fourth International Workshop on Genetic and Evolutionary Fuzzy Systems, Aristotle University of Thessaloniki, Mieres, Spain, 2010.

Daza G. S, Sánchez I. G, Suárez J. C, Selección de Características Orientada a Sistemas de Reconocimiento de Granos Maduros de Café, Scientia et Técnica, año XIII No 35, Universidad Tecnológica de Pereira, 2007, pp 139-144.

Ekenel H, Sankur B, Feature selection in the independent component subspace for face recognition, Pattern Recognition Letters 25 (2004) 1377–1388, ELSEVIER.

Fuentes G, et. al., Wavelet Filters Applied to Automatic Face Recognition, GEST Int'l Trans. Communication and Signal Processing, 2006, Vol. 8, No1.

Fuentes Gibran, Reconocimiento de Rostros Mediante Wavelets y Redes Neuronales, Tesis presentada al Instituto Politécnico Nacional, México, para optar al título de magister, 2008.

Gómez, L, et. al., Semi-supervised method for crop classification using hyperspectral remote sensing images, 1st International Symposium. Recent Advantages in Quantitative Remote Sensing, RAQRS'2002. pp. 488–495. Torrent (Valencia).

Grefenstette J, Optimization of control parameters for genetic algorithms, IEEE Transactions on Systems, Man and Cybernetics, Vol. 16, N° 1, 1986, pp. 122-128.

J.A. Cohen, Coefficient of agreement for nominal scales. Educational and Psychological Measurement, 1960, pp. 37–46.

Jeong D, et. al., Feature Selection for Steel Defects Classification, International Conference on Control, Automation and Systems 2010, Oct. 27-30, 2010 in KINTEX, Gyeonggi-do, Korea, 2010.

Landis, J & Koch, G. The measurement of observer agreement for categorical data. Biometrics 33, 1977, pp. 159-174.

Liu J, Wang G, A Hybrid Feature Selection Method for Data Sets of thousands of Variables, 978-1-4244-5848-6/10/ ©2010 IEEE.

Mitra S, Sankar K. P, Fuzzy sets in pattern recognition and machine intelligence, Fuzzy Sets and Systems, vol. 156, 2005, pp 381–386.

Orozco, M, Selección Efectiva De Características Wavelet En La Identificación De Bioseñales 1-D Y 2-D Usando Algoritmos Genéticos, Tesis presentada a Universidad Nacional de Colombia para optar al grado de Magister, 2005.

Pazoki Z, Farokhi F, Effective feature selection for face recognition based on correspondence analysis and trained artificial neural network, Sixth International Conference on Signal-Image Technology and Internet Based Systems, 2010.

Quintero C, Gamarra M, Niño J, Control de Calidad Inteligente de Piezas Automotrices Basado en Lógica Difusa Mediante Procesamiento Digital de Imágenes, en Memorias XV Simposio de Tratamiento de Señales, Imágenes y Visión Artificial, pp. 150-156, ISBN 978-958-8060-96-5, Bogotá, Colombia, 2010.

Rivera J.H, Castellanos C, Soto J.M, Selección Efectiva de Características para Bioseñales Utilizando el Análisis de Componentes Principales, Scientia et Técnica, año XIII No 34, Universidad Tecnológica de Pereira, 2007, pp. 127-131.

Song F, Guo Z, Feature selection using principal component analysis, 2010 International Conference on System Science, Engineering Design and Manufacturing Informatization, 2010.

Theodoridis S, Koutroumbas K, Pattern Recognition, Fourth Edition, Canada, 2009.

Umamaheswari, K, Sivanandam S.N, Neuro – Genetic Approaches To Classification Of Face Images With Effective Feature Selection Using Hybrid Classifiers, PSG College of Technology, Coimbatore, India, 2006.

Vieira S, et. al., Cohen's Kappa Coefficient as a Performance Measure for Feature Selection, IEEE, 2010.

Vilches E, Data Mining Applied to Acoustic Bird Species Recognition, 18th International Conference of Pattern Recognition 2006 (ICPR 06), Hong Kong.

Wang Y, Fan K, Applying Genetic Algorithms on Pattern Recognition: An Analysis and Survey, Proceedings of ICPR '96, IEEE, 1996.