



## XVII<sup>th</sup> World Congress of the International Commission of Agricultural and Biosystems Engineering (CIGR)

Hosted by the Canadian Society for Bioengineering (CSBE/SCGAB)  
Québec City, Canada June 13-17, 2010



### CLASSIFICATION OF PRE-SLICED HAM IMAGES WITH QUATERNIONIC SINGULAR VALUES USING AN ADAPTIVE MULTILAYER PERCEPTRON

NEKTARIOS A. VALOUS<sup>1</sup>, DA-WEN SUN<sup>1</sup>

<sup>1</sup>N.A. VALOUS, FRCFT Group, Biosystems Engineering, Agriculture & Food Science Centre, School of Agriculture Food Science & Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland.

<sup>1</sup>D.-W. SUN, Tel: +353 1 716 7342, Fax: +353 1 716 7493, dawen.sun@ucd.ie, www.ucd.ie/refrig.

#### CSBE100264 – Presented at Section VI: Postharvest Technology and Process Engineering Conference

**ABSTRACT** The quaternionic representation of ham images, treating RGB colour components as a single unit instead of as separate components, is very effective. The advantage of using quaternion arithmetic is that the perceptually richer colour images can be represented and analyzed as a single entity, improving the accuracy of pattern recognition models. The quaternionic singular value decomposition (SVD) is a technique to decompose a quaternion matrix into quaternion singular vector and singular value component matrices exposing useful properties. Singular values describe completely and univocally the intrinsic information of a quaternionic matrix, ergo they can be used as features for the classification of pork ham slices. The objective was to use a small portion of uncorrelated singular values, as robust features for the classification of sliced ham images, using a supervised artificial neural network classifier. Images were acquired from four qualities of sliced cooked pork ham typically consumed in Ireland (90 slices/quality), having similar appearances. Mahalanobis distances and Pearson product moment correlations were used for feature selection. The dimensionality reduction procedure excluded atypical features and discarded the redundant information. An adaptive multilayer perceptron classifier was successfully employed, using a reduced feature space of six singular values. The overall correct classification performance for the test set was 86.1%. Results confirmed that the classification performance was satisfactory. Using the most informative features as input to the multilayer perceptron classifier led to the recognition of a set of different but visually quite similar textural patterns.

**Keywords:** Computer vision, Pork ham slice, Supervised classification, Quaternionic singular value decomposition, Quaternionic singular values, Mahalanobis distance, Artificial neural network, Multilayer perceptron.

**INTRODUCTION** Although gray level images can be quite satisfactory from the pattern recognition perspective, RGB colour images alternatively do seem to be perceptually richer. A quite recent approach is to encode the three channel components (R, G, and B) on the three imaginary parts of a quaternion (Denis et al., 2007). There is a growing interest in the applications of quaternion numbers to colour image processing. Many problems in the area of quaternion-based image processing are still open (Cai & Mitra, 2000). In general, quaternions are an extension of complex numbers to four dimensions

and play an important role in colour image processing. They are considered as complex numbers with a vector imaginary part consisting of three mutually orthogonal components. In Cartesian form, quaternion numbers can be represented as follows:

$$quat = a_1 + o_1a_2 + o_2a_3 + o_3a_4 \quad (1)$$

where  $a_1, a_2, a_3$  and  $a_4$  are real numbers, and  $o_1, o_2$  and  $o_3$  are orthogonal imaginary operators. A pure quaternion has a zero real part and a full quaternion has a non-zero real part. Thus, a quaternion number is a complex number with real and imaginary parts, hence the term hypercomplex (Kantor & Solodovnikov, 1989; Sangwine, 1996). An RGB colour image of size  $(x,y)$  may be converted to a quaternion matrix by placing the three colour components into the three quaternionic imaginary parts, leaving the real part zero such that the image function  $A_q(x,y)$  is given by the following representation (Moxey et al., 2003):

$$A_q(x, y) = A_R(x, y)i + A_G(x, y)j + A_B(x, y)k \quad (2)$$

where  $A_R(x,y)$ ,  $A_G(x,y)$ , and  $A_B(x,y)$  are the R, G, and B components, respectively, at the pixel coordinates of the image. In this way, the colour image is represented as a matrix of size  $[x \cdot y]$  whose elements are pure quaternions (Ell & Sangwine, 2007). Although real and complex number systems are used to provide arithmetic operations of 1D and 2D data, quaternions can handle algebraic operations of ternary numbers, ergo expressing colour data directly (Pei & Cheng, 1997). The quaternionic singular value decomposition (SVD) is a technique to decompose a quaternion matrix (representation of a colour image) into several component matrices, exposing useful properties of the original matrix. SVD has been exploited generally in what is known as reduced-rank signal processing where the idea is to extract the significant parts of a signal (Sangwine & Le Bihan, 2006). In image pattern recognition, identifying and extracting effective features is an important step to successfully complete the task of classification. There are several kinds of image features for recognition such as visual, algebraic, statistical moments and transform coefficients (Hong, 1991). Algebraic features represent intrinsic attributions of the image, ergo various transforms or decompositions can be used to extract them. The quaternionic SVD is an effective algebraic feature extraction method for any colour image (Hong, 1991). Since the quaternionic matrix designates a mathematical representation of a colour image, the computed singular value feature vectors are unique for the colour image and can be used for pattern recognition purposes. SVD has many generalizations and refinements depending on the broad range of its applications in science and engineering. The SVD of the ham image, expressed with the quaternion arithmetic, produces unique singular values (descriptors) that are identified as algebraic features to potentially recognize/classify ham colour images robustly.

In the context of classification, neural networks can be viewed as powerful, fault tolerant and reasonable alternatives to traditional classifiers. Multilayer perceptron (MLP) networks with one hidden layer and sigmoidal hidden layer activation functions are capable of approximating any decision boundary to arbitrary accuracy (Li et al., 1999), ergo could be employed to learn a given mapping from a finite singular value dataset. MLPs are used due to their popularity and enhanced ability for generalization, which is related to the accurate prediction of data that are not part of the training dataset. A

comprehensive description of the salient features of an MLP neural network can be found in Bishop (1995). The objective of this study is to classify four qualities of cooked pork ham typically consumed in Ireland, with a supervised MLP neural network, using a small portion of informative and uncorrelated singular values computed from the quaternionic SVD of digital colour images, as robust and stable features. To our knowledge this is the first reported use of quaternions and quaternionic singular values in food image analysis and in particular for the classification and quality grading of images of food surfaces.

## MATERIALS AND METHODS

**Pork ham samples** Four cooked pork ham qualities were manufactured in the agriculture and food development authority Teagasc (Co. Dublin, Ireland) using Silverside pork leg muscles (*M. biceps femoris*) without membranes and sinew ends. The muscles were injected with different percentages of brine solutions (wet curing by injection). The resulting product qualities were: premier quality or low yield ham (A1; 10% injection level), medium quality hams (A2 and A3 with 20% and 30% injection levels, respectively) and low quality or high yield ham (A4; 40% level). The averaged moisture content of ham slices in % per quality along with the standard deviation was the following: A1 (71.8  $\pm$ 0.4), A2 (72.6  $\pm$ 0.5), A3 (74.9  $\pm$ 0.6), and A4 (74.8  $\pm$ 0.5). The moisture content was determined in quadruplicate by the official AOAC method 950.46, which is related to the determination of moisture content in meat and meat products (AOAC, 1998). Pre-forming tumbling was carried out for 30 minutes at 6 rpm for all hams. The injected muscles were vacuum tumbled for 6 up to 20 hours depending on the quality. The injected and tumbled pork muscles were formed, vacuum packed and pressed into shape using pressure moulds before steam cooking. The hams were cooked at 82°C to a core temperature of 74°C. All pork ham samples were chilled to 4°C before slicing (slice width  $\approx$  2.0 mm). Images were acquired immediately after slicing (90 slices per quality).

**Image acquisition and processing** A colour calibrated computer vision system (CVS) as described by Valous et al. (2009) was used for image acquisition (spatial resolution of 0.0826 mm/pixel). The software package MATLAB v7.4 (MathWorks, USA) along with the open-source quaternion toolbox (Sangwine & Le Bihan, 2005) was used for image processing and singular value extraction. The toolbox allows computations with quaternion matrices in almost the same way as with matrices of complex numbers. A polynomial transform for calibrating colour signals (Valous et al., 2009) was used to map the RGB primaries to sRGB, ensuring reproducible colour images. Due to the differences in size and shape among the four qualities and the considerable computational delays associated with the quaternionic SVD, the acquired colour images were subsequently cropped in the central region (1024x1024 pixels; equivalent to 7154.1 mm<sup>2</sup>) to produce sixteen 256x256 pixel sub-images (equivalent to 447.1 mm<sup>2</sup>) per ham slice image and per quality. Cropping allowed better scrutiny and interpretation (all analyzed images had the same spatial pixel dimensions) and also kept computation times manageable. Thus, the colour calibrated sub-images, expressed as quaternion matrices, were the direct input to the quaternionic SVD algorithm for the extraction of singular values.

**Singular value extraction from quaternionic SVD** In general, SVD is related to the diagonalization of a matrix and is a well-used tool in the theory of linear algebra. Without separating the image into three separate colour channel images, it has been demonstrated

that this colour image decomposition exists based on quaternion matrix algebra. Specifically, in quaternionic SVD, for any arbitrary  $x \cdot y$  quaternion matrix  $A_q$ , there exist two quaternion unitary matrices  $U$  and  $V$  and a diagonal  $\Sigma$  such that the following factorization exists in the following form (Le Bihan & Mars, 2004):

$$A_q = U\Sigma V^T, \text{ with } \Sigma = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} \quad (3)$$

where  $U$  denotes an  $x \cdot x$  unitary quaternion matrix,  $\Sigma$  is an  $x \cdot y$  diagonal matrix with non-negative real numbers on the diagonal, and  $V^T$  is the conjugate transpose of an  $x \cdot y$  unitary quaternion matrix  $V$ . The diagonal entries of  $\Sigma$ , called singular values  $\Sigma_r$ , can be arranged in order of decreasing magnitude and the columns of  $U$  and  $V$  are called left and right quaternion singular vectors for  $A_q$ , respectively. The quaternionic SVD can be rewritten as singular value spectrum decomposition; the matrix  $A_q$  with rank  $r$  has a Fourier expansion, in terms of the singular values and outer products of the columns of the  $U$  and  $V$  matrices of the following form (Gentle, 2007):

$$A_q = \sum_{n=1}^r u_n v_n^T \sigma_n \quad (4)$$

where  $u_n$  are the left singular vectors (columns of  $U$ ),  $v_n^T$  the right singular vectors (columns of  $V$ ), and  $\sigma_r$  are the real singular values. Equation (4) shows that this quaternionic factorization, decomposes the matrix into a sum of  $r$  rank-1 quaternion matrices, ergo it is said to be rank revealing; the number of non-null singular values equals the rank of the matrix (Le Bihan & Sangwine, 2007). The method for obtaining the SVD in colour images expressed as quaternion matrices can be found in Le Bihan and Mars (2004). From the algorithmic development perspective, deriving the SVD using the Jacobi algorithm produces more accurate results than any other known algorithm (Le Bihan & Sangwine, 2007). A much faster algorithm based on the transformation of a quaternion matrix to bidiagonal form, using quaternionic Householder transformations has been developed in a previous study (Sangwine & Le Bihan, 2006). This algorithm is less accurate than the Jacobi algorithm, but has significant computational speed advantages. Moreover, it is the default quaternionic SVD algorithm that has been implemented in the quaternion MATLAB toolbox.

**Feature space reduction** An operation such as a classification that would have been performed on the colour image can now be equivalently performed on the real non-negative diagonal elements  $\sigma_r$ . Thus, the 256 singular values computed for each of the sixteen sub-images (forming the central region of each ham slice) were linearly rescaled (Barcala et al., 2004), using min-max normalization (values from 0 to 1) and averaged to obtain an array consisting of 256 features, representing each of the 90 ham slices. The final result was a matrix of 256 pre-processed (averaged and normalized) features (F1, F2, F3 ... F256) per image and per quality. In this way, the pork ham slice was sampled along vertical and horizontal dimensions yielding a set of sixteen sub-images, from which their averaged singular values represented the initial square image. The average coefficients of variation (expressed as percentages), which describe the standard deviations as a percentage of the averaged singular values computed from the sub-images, showed that there were not any substantial variations within the square slice

images (1024 x 1024 pixels). Mahalanobis distance and Pearson product moment correlations were used as dimensionality reduction tools for feature selection. Mahalanobis distance is a distance metric based on feature correlations among ham qualities and is useful as a means to determine similarity, between pairs of groups (Mendoza et al., 2009). Specifically, the Mahalanobis distance computation in MATLAB, among singular values was carried out across six different ham quality pairs (A1-A2, A1-A3, A1-A4, A2-A3, A2-A4, and A3-A4) as follows (Tay et al., 2007):

$$D(f_1, f_2) = (f_1 - f_2)^T \text{Cov}(f_1, f_2)(f_1 - f_2) \quad (5)$$

where  $D(f_1, f_2)$  is the Mahalanobis distance,  $f_1$  and  $f_2$  are the set of singular values for any two ham quality pairs, and given the matrix  $m_a \times d$  of the dataset ( $m_a$  data vectors of dimensionality  $d = 256$ ), the covariance of the  $i$ th feature  $x_i$  ( $\mu_i$ : corresponding mean) and the  $j$ th feature  $x_j$  ( $\mu_j$ : corresponding mean) are computed using the following:

$$\text{Cov}(x_i, x_j) = \frac{1}{n} \sum_{k=1}^n (x_{ki} - \mu_i)(x_{kj} - \mu_j) \quad (6)$$

Eighteen features (F1, F2, F5, F7, F8, F9, F10, F11, F12, F13, F14, F15, F18, F21, F89, F245, F251 and F252) were selected. This was achieved by choosing (among the 256) the features that exhibited the five largest distances between each pair of ham qualities (denoting maximum separability). This metric was used as an initial scrutiny tool to reduce data dimensionality (feature set S1 = 18). However, some of the selected features are highly correlated, so Pearson product moment correlation coefficients ( $r$ ) were computed to measure the strength of the linear relationships among the previously selected feature set. This procedure defined a mapping from a typically higher dimensional data space to a space of reduced dimension (feature set S2 = 6; F1, F2, F5, F7, F89 and F252), maintaining key data properties and resulting in a less correlated dataset in the range of correlations -0.6 to 0.6. Hence, only a handful of potential highly discriminating features were used as input to train the neural network.

**Supervised artificial neural network classification** A supervised MLP neural network was employed for classification. The MLP classifier consists of a set of simple processing units arranged in a layered architecture. MLP is concerned into partitioning the feature space and finding class memberships, determined by the categorical levels of four ham qualities for the given input set S2 (6 features) of  $\sigma_r$ , with the target to lead to high classification rates. Using the software package STATISTICA 8.0 (StatSoft, USA), the weights connecting the inputs to the hidden neurons and the hidden neurons to the output neurons were adjusted, so that the network could be trained by approximating the underlying functional relationship between the training set (60% randomly selected data; 54 images per quality) and the target ham classes. Given that the size of the training set can have a significant effect on classification accuracy more data were used for training than for validation and testing. The cross entropy penalty (error) function, which is more suitable for classification problems, evaluated the performance of the MLP during training, measuring how close the network predictions were to the targets and how much weight adjustment should be applied by the algorithm in each iteration (Fontenla-Romero et al., 2005). To assess the performance while under training, a randomly selected unseen validation dataset (20%; 18 images per quality) was chosen as a means of checking how

well the network makes progress in modelling the input-target relationship. Such an assessment (cross-validation) is necessary to avoid the overtraining (overfitting) phenomenon, which causes displacement of decision boundaries. Regularization, using weight decay, was also considered for improving the generalization of the MLP, adding a term to the error function which penalizes large weight values. This form of regularization can lead to significant improvements in network generalization (Lerouge & Van Huffel, 1999). The remaining images (18 images per quality) were used as the test set to measure the classification accuracy of the neural network on unseen data. The data selection process for the training, validation and test set was carried out using MATLAB's random integer number generator function 'randi'.

**RESULTS AND DISCUSSION** Due to the lack of a concrete rule for choosing the optimum number of neurons for the hidden layer, preliminary trial and error tests were carried out to determine the number of hidden neurons in order to build the neural classifier. In general, the more neurons the hidden layer contains the more flexible it becomes, increasing the classification accuracy for the training data but decreasing the accuracy for the test set (Bishop, 1995). More specifically, after a certain threshold of neurons has been reached, increasing their number beyond that threshold has a marginal effect on the resulting performance of the classifier (Teoh et al., 2006). In these tests, a small number of hidden layer nodes (2 - 5) produced higher training and high generalization error due to underfitting and high statistical bias, while larger number of hidden nodes (7 - 20) increased the training classification performance, but produced a higher generalization error due to overfitting and high variance. Consequently, the selection of the neural network architecture was based on reaching a compromise between too many and too few neurons in the hidden layer. The best generalizing neural network is not necessarily the one with the fewest number of hidden neurons (Kinser, 2000). In addition, the number of hidden units determines the total number of weights in the network and thus there should not be more weights than the total number of training points in the dataset (Duda et al., 2001). Figure 1 shows a schematic of the classifier architecture used.

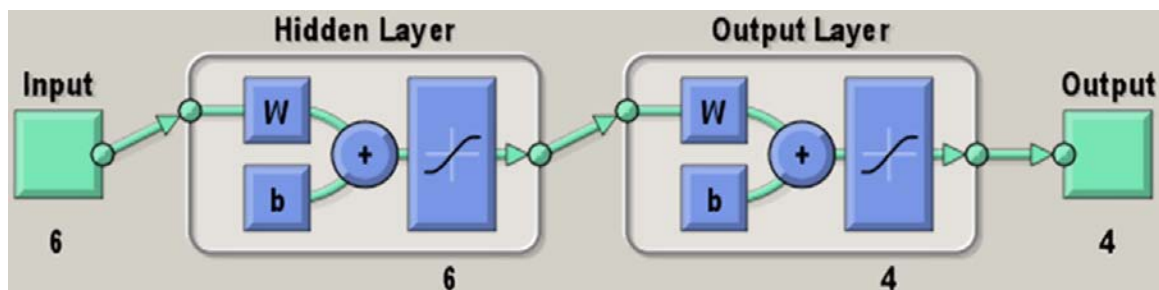


Figure 1. Schematic of the MLP neural network classifier architecture (6-6-4) used for pattern recognition (W are weights and b are biases).

The optimal conditions for the classification were found to be a single hidden layer composed of 6 neurons (6-6-4 architecture; 6 features as input) having symmetric sigmoid (hyperbolic tangent) activation functions, and 4 output neurons corresponding to ham qualities having softmax activation functions, in which the outputs are interpretable as posterior probabilities for the categorical target variables. Using softmax activation functions in the output layer provides a measure of certainty, while classification

accuracy is improved (Dunne, 2007). An adaptive supervised feedforward multilayer perceptron classifier, using a variant of the Quasi-Newton method; namely the BFGS/BP (Broyden Fletcher Goldfarb Shanno/Back Propagation) learning algorithm, was employed to obtain a suitable mapping from the input dataset. This algorithm performs better in terms of training speed and accuracy and requires more computation in each iteration and more storage, but has a fast convergence rate. The network was trained for 58 epochs (passes through the entire training set). A convenient rule of thumb is that when both validation and test datasets produce good and consistent classification results, it could be assumed that the network generalizes well on unseen data. The results of the classification are presented in Table 1 as confusion matrices. The overall correct classification performance for the training, validation and test set of singular values were 90.3%, 94.4% and 86.1%, respectively.

Table 1. Neural classifier output in the form of confusion matrices for the training, validation and test set of quaternionic singular values.

<b>TRAINING SET</b>					
<b>Predicted</b>	<b>A1</b>	<b>A2</b>	<b>A3</b>	<b>A4</b>	<b>Overall classification performance</b>
<b>A1</b>	50 <sup>++</sup>	3	8	4	
<b>A2</b>	1	50 <sup>++</sup>	0	0	
<b>A3</b>	3	1	46	1	
<b>A4</b>	0	0	0	49 <sup>++</sup>	
<b>Image set</b>	54	54	54	54	90.3%
<b>VALIDATION SET</b>					
<b>Predicted</b>	<b>A1</b>	<b>A2</b>	<b>A3</b>	<b>A4</b>	<b>Overall classification performance</b>
<b>A1</b>	17 <sup>++</sup>	0	1	1	
<b>A2</b>	0	18 <sup>++</sup>	0	0	
<b>A3</b>	1	0	17 <sup>++</sup>	1	
<b>A4</b>	0	0	0	16	
<b>Image set</b>	18	18	18	18	94.4%
<b>TEST SET</b>					
<b>Predicted</b>	<b>A1</b>	<b>A2</b>	<b>A3</b>	<b>A4</b>	<b>Overall classification performance</b>
<b>A1</b>	14	0	1	0	
<b>A2</b>	0	16	0	0	
<b>A3</b>	4	2	15	1	
<b>A4</b>	0	0	2	17 <sup>++</sup>	
<b>Image set</b>	18	18	18	18	86.1%

++, denotes classification errors less than  $\leq 10\%$ .

The validation and test (generalization) errors are virtually always higher than the training error (Duda et al., 2001). From the results, it can be seen that the generalization capability of the neural network is satisfactory for the relatively small dataset of 90 images per quality. The training inputs to the neural classifier contained sufficient information pertaining to the target qualities. In spite of the high variability and

complexity of the studied ham samples, the results showed the capacity of  $\sigma_7$  to provide valuable information in discerning among different qualities. Using a bigger number of uncorrelated features most probably would increase the classification accuracy for the test set. Nevertheless, preliminary classification tests using the ten largest (in magnitude) singular values ( $\sigma_1 - \sigma_{10}$ ) as input features to the neural classifier showed that the recognition errors were greater than those presented in Table 1. This could be attributed most likely to the fact that  $\sigma_3, \sigma_4, \sigma_6, \sigma_8, \sigma_9$  and  $\sigma_{10}$  are highly correlated among themselves and with the already selected uncorrelated features  $\sigma_1, \sigma_2, \sigma_5$ , and  $\sigma_7$ , thus lowering the classification performance. In relation to this deduction, the principle of parsimony states that the smallest possible number of features should be used so as to give an adequate and uncorrelated representation of the feature space (Chatfield, 1996). The results indicate that the classifier performs well on unseen data (test set). This deduction is based on the experimental results of the MLP-based neural network; therefore it cannot be generalized to other classification techniques.

The analysis of confusion matrices for the test set yields interesting insights into the differences in texture appearance and the global properties of the neural classifier. Due to the fact that pork hams are coming from muscles that exhibit normal biological variations and are subject to industrial processing/storage conditions that are not exactly replicated, it is unrealistic to expect 100% accurate classification at all times. Therefore, it will be necessary to set a cut-off for acceptance of a correctly classified ham slice. More than 80-85% consistent classification rate responding to new samples can be such a threshold, given that a good artificial neural network model has accuracy of more than 80% (Hanif et al., 2009). The tendency of the misclassification rate is to decrease towards the lower quality ham, with the exception of an increase in A3. Results also show that the A1 ham in the test set had the worse classification performance (classification error; 22%) comparing with the error in the other qualities (16%). This erroneous assignment to a class membership (A3) other than the correct one could be probably attributed to the lack of a sufficiently diverse training set of singular values for this ham quality. Another interesting deduction is that the driest samples (A1) were more difficult to classify. Even when the generalization for this quality was weak (less than 80%), there is still a significant dependency between  $\sigma_7$  and neural network output due to the results in the training and validation set. A3 apparently shared some of the underlying visual texture characteristics of the A1 quality and the reduced generalization performance of the classification for this case seems to suggest the hypothesis that singular values preserve important topological properties of the input images. On the other hand, the lowest quality A4 ham produced singular values that had optimal discriminating properties, with only one misclassified image. The A4 pork ham was manufactured with the highest level of brine injection and an increased duration of tumbling, which resulted in wetter surface appearance and an intermediate degree of visual roughness. Regardless of this textural complexity, singular values captured relevant information that provided a good level of differentiation.

**CONCLUSION** Simple visual features such as colour and texture, as well as spatial features such as shape and distribution of structures (fat-connective tissue and pores/defects) contribute to the complexity of texture appearance. The quaternionic representation of ham images, treating RGB colour components as a single unit instead of as separate components, is very effective. The advantage of using quaternion arithmetic is that colour images, which are perceptually richer, can be represented and analyzed as a



single entity, improving the accuracy of pattern recognition models. Algebraic features represent intrinsic attributions of an image. The quaternionic SVD is an effective method of extracting algebraic features from ham images. Singular values describe completely and univocally the intrinsic information of a quaternionic matrix, ergo they can be used as features for the classification of cooked pork ham slices. An adaptive MLP classifier was successfully employed for the classification of four ham qualities with similar appearances, using a reduced feature space of singular values. The dimensionality reduction procedure excluded atypical features and discarded the redundant information. The overall correct classification performance for the test set was 86.1%. Considering the complexity of texture appearance, it is difficult to get perfect classification rates using neural networks based on certain selected features. Nonetheless, accurately extracting and selecting the most informative features as inputs to the MLP classifier, led to the recognition of a set of different but visually quite similar textural patterns based on quaternionic singular values.

**Acknowledgements** The authors gratefully acknowledge the Food Institutional Research Measure (FIRM) strategic research initiative, as administered by the Irish Department of Agriculture, Fisheries & Food, for the financial support.

## REFERENCES

- Barcala, J.M., Fernández, J.L., Alberdi, J., Jiménez, J., Lázaro, J.C., Navarrete, J.J., & Oller, J.C. 2004. Identification of plastics using wavelets and quaternion numbers. *Measurement Science & Technology*, 15, 371-376.
- Bishop, C.M. 1995. *Neural Networks for Pattern Recognition* (pp. 116-164). Oxford: Clarendon Press.
- Cai, C., & Mitra, S.K. 2000. A normalized color difference edge detector based on quaternion representation. In *Proceedings of the IEEE International Conference on Image Processing* (pp. 816-819), 10-13 September 2000, Vancouver, Canada.
- Chatfield, C. 1996. Model uncertainty and forecast accuracy. *Journal of Forecasting*, 15(7), 495-508.
- Denis, P., Carre, P., & Fernandez-Maloigne, C. 2007. Spatial and spectral quaternionic approaches for colour images. *Computer Vision and Image Understanding*, 107, 74-87.
- Duda, R.O., Hart, P.E., & Stork, D.G. 2001. *Pattern Classification*, 2nd edition (pp. 282-349). New York: John Wiley & Sons.
- Dunne, R.A. 2007. *A Statistical Approach to Neural Networks for Pattern Recognition* (pp. 35-51). New Jersey: John Wiley & Sons.
- Ell, T.A., & Sangwine, S.J. 2007. Hypercomplex Fourier transforms of color images. *IEEE Transactions on Image Processing*, 16(1), 22-35.
- Fontenla-Romero, O., Guijarro-Berdiñas, B., Alonso-Betanzos, A., & Moret-Bonillo, V. 2005. A new method for sleep apnea classification using wavelets and feedforward neural networks. *Artificial Intelligence in Medicine*, 34(1), 65-76.
- Gentle, J.E. 2007. *Matrix Algebra: Theory, Computations, and Applications in Statistics* (pp. 41-144). New York: Springer.
- Hanif, N.H.H.M., Lan, W.H., Daud, H.B., & Ahmad, J. 2009. Classification of control measures for asthma using artificial neural networks. In *Proceedings of the 9th IASTED International Conference on Artificial Intelligence and Applications* (639-069), 17-18 February 2009, Innsbruck, Austria.
- Hong, Z-Q. 1991. Algebraic features extraction of image for recognition. *Pattern Recognition*, 24(3), 211-219.

- Kantor, I.L., & Solodovnikov, A.S. 1989. *Hypercomplex Numbers: An Elementary Introduction to Algebras* (pp. 15-24). New York: Springer-Verlag.
- Kinser, J.M. 2000. The minimum number of hidden neurons does not necessarily provide the best generalization. In *Proceedings of SPIE Applications and Science of Computational Intelligence III* (vol. 4055, pp. 11-17), 24-27 April 2000, Orlando, USA.
- Le Bihan, N., & Mars, J. 2004. Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing. *Signal Processing*, 84, 1177-1199.
- Le Bihan, N., & Sangwine, S.J. 2007. Jacobi method for quaternion matrix singular value decomposition. *Applied Mathematics and Computation*, 187, 1265-1271.
- Lerouge, E., & Van Huffel S. 1999. Generalization capacity of neural networks for the classification of ovarium tumors. In *Proceedings of the 20th Symposium on Information Theory in the Benelux* (pp. 149-156), 27-28 May 1999, Haasrode, Belgium.
- Li, Y., Rad, A.B., & Peng, W. 1999. An enhanced training algorithm for multilayer neural networks based on reference output of hidden layer. *Neural Computing and Applications*, 8, 218-225.
- Mendoza, F., Valous, N.A., Allen, P., Kenny, T.A., Ward, P., & Sun, D.-W. 2009. Analysis and classification of commercial ham slice images using directional fractal dimension features. *Meat Science*, 81(2), 313-320.
- Moxey, C.E., Sangwine, S.J., & Ell, T.A. 2003. Hypercomplex correlation techniques for vector images. *IEEE Transactions on Signal Processing*, 51(7), 1941-1953.
- Pei, S-C., & Cheng, C-M. 1997. A novel block truncation coding of color images using a quaternion-moment-preserving principle. *IEEE Transactions on Communications*, 45(5), 583-595.
- Sangwine, S.J. 1996. Fourier transforms of colour images using quaternion or hypercomplex, numbers. *Electronic Letters*, 32(21), 1979-1980.
- Sangwine, S.J., & Le Bihan, N. 2005. Quaternion Toolbox for MATLAB. Available at: <http://qtfm.sourceforge.net/>. Software library, licensed under the GNU General Public License. Accessed 11 February 2009.
- Sangwine, S.J., & Le Bihan, N. 2006. Quaternion singular value decomposition based on bidiagonalization to a real or complex matrix using quaternion Householder transformations. *Applied Mathematics and Computation*, 182, 727-738.
- Tay, A.L.P., Zurada, J.M., Wong, L-P., & Xu, J. 2007. The hierarchical fast learning artificial neural network (HieFLANN) - an autonomous platform for hierarchical neural network construction. *IEEE Transactions on Neural Networks*, 18(6), 1645-1657.
- Teoh, E.J., Tan, K.C., & Xiang, C. 2006. Estimating the number of hidden neurons in a feedforward network using the singular value decomposition. *IEEE Transactions on Neural Networks*, 17(6), 1623-1629.
- Valous, N.A., Mendoza, F., Sun, D.-W., & Allen, P. 2009. Colour calibration of a laboratory computer vision system for quality evaluation of pre-sliced hams. *Meat Science*, 81(1), 132-141.