

MORPHOLOGICAL GRANULOMETRY FOR CLASSIFICATION OF EVOLVING AND ORDERED TEXTURE IMAGES

Mahmuda Khatun⁽¹⁾, Alison Gray⁽¹⁾, and Stephen Marshall⁽²⁾

⁽¹⁾ Department of Mathematics and Statistics,
University of Strathclyde, Glasgow G1 1XH, UK
phone: +44 (0)141 548 4335, fax: +44(0)141 548 3345,
email: m.khatun, a.j.gray@strath.ac.uk

⁽²⁾ Department of Electronic and Electrical Engineering,
University of Strathclyde, Glasgow, G1 1XW, UK
phone: +44 (0)141 548 2199, fax: +44(0) 141 552 2487,
email: s.marshall@eee.strath.ac.uk

ABSTRACT

In this work we investigate the use of morphological granulometric moments as texture descriptors to predict time or class of texture images which evolve over time or follow an intrinsic ordering of textures. A cubic polynomial regression was used to model each of several granulometric moments as a function of time or class. These models are then combined and used to predict time or class. The methodology was developed on synthetic images of evolving textures and then successfully applied to classify a sequence of corrosion images to a point on an evolution time scale. Classification performance of the new regression approach is compared to that of linear discriminant analysis, neural networks and support vector machines. We also apply our method to images of black tea leaves, which are ordered according to granule size, and very high classification accuracy was attained compared to existing published results for these images. It was also found that granulometric moments provide much improved classification compared to grey level co-occurrence features for shape-based texture images.

1. INTRODUCTION

Texture classification is a long-standing problem in image processing. It involves a feature extraction step, in which a set of texture features is extracted from the image under study, and a classification step, in which a texture class membership is assigned to it based on the information provided by the extracted texture features through appropriate machine learning algorithms [12]. Here we use morphological granulometry and co-occurrence matrix approaches to extract texture features, and employ a regression-based texture classification approach to compare the relative usefulness of the two different sets of features for classifying sequences of texture images which either evolve over time or follow an intrinsic ordering. For example, size of texture primitives may increase over time or with class label, as in spots of corrosion building up on sheet metal.

Morphological granulometry [13] is extensively used to extract textural information from images. It was first introduced to characterise size and shape information for a binary image, considered to be a collection of grains. The concept of granulometry is to sieve the grains through filters of increasing size, so that grains with size smaller than the holes will drop out and only grains with larger size will remain. The shape of the holes is determined by the shape of the *structuring element* (SE), which is a geometrical pattern used to extract textural information from a given digital image [6]. In the sieving, the remaining image area successively decreases

and eventually drops to zero. The rate of decrease represents the cumulative proportion of image area dropped, known as the *size distribution* [1]. A probability distribution function (pdf) can be derived from this size distribution, since it is a cumulative distribution function (cdf). Its statistical moments, known as *granulometric moments*, contain useful textural information to characterise the pdf and the image.

Texture features derived from the grey level co-occurrence matrix (GLCM) [9] are the most widely used texture features for classification. The GLCM represents the probability distribution of occurrence of a pixel pair, at a given separation and at a given orientation, with given grey levels. Various texture features can be derived from the GLCM. GLCM features were successfully used by Chanda and Majumder [3] to segment images of chromosomes. Soh and Tsatsoulis [16] obtained 94.17% classification accuracy for SAR sea ice images using GLCM features in Bayesian classifiers. Clausi [4] computed 8 different GLCM features using different quantisations of SAR sea ice images, and used the features jointly and separately to classify the images.

In this work a regression-based classifier is developed by modelling granulometric moments or GLCM features as a function of texture evolution time or class label. A cubic polynomial regression is fitted for each chosen feature separately, and then a combined cubic polynomial regression is obtained. The combined model is used for back-prediction of evolution time or class label of a new image using its observed texture features. Linear discriminant analysis (LDA), neural networks and support vector machines (SVMs) are used with the same sets of features (either granulometric moments or GLCM features) to compare their classification accuracy with that of the new regression approach.

We are interested to compare the usefulness of granulometric moments and GLCM features in discriminating texture images where textures represent some sort of damage or decay which progresses over time. Knowing the state of damage is vital in many cases, for example, regular monitoring of degree of damage of machine parts is of crucial importance in industrial inspection. Classification of images of corroding metal according to their degree of corrosion was studied in [14, 15, 7], and the same set of corrosion images are used here as an example of textures that evolve over time.

Another application is the sorting of tea into different grades according to granule size. This is very important task in the tea processing industry, which has traditionally been carried out by sieving with a series of sieves of differently sized mesh. More recently, computer vision and pattern recognition have been investigated for a more automated ap-

proach [2]. We have also applied our methodology to a sequence of black tea images representing different grades of tea of different granule sizes [2]. This is an example of textures ordered by class.

2. METHODOLOGY

Here we briefly describe the feature extraction methods and different classifiers used.

2.1 Morphological granulometry

Morphological techniques are widely used in digital image processing. The foundation of morphological processing is in set theory. One of the fundamental techniques, i.e. *opening*, provides very useful information about shape and size of image objects or texture primitives. Opening of an image f by a SE g , denoted by $f \circ g$, is the union of all translations of g that are a subset of f . The effect of opening can be explained as sliding SE g beneath the input image f , eliminating any details smaller than g and reducing the height or grey level of larger objects.

Granulometry consists of successive openings of an image by a sequence of SEs of increasing size. An image f is opened sequentially by a series of scaled SEs $\{g_1, g_2, \dots, g_N\}$, e.g. successively larger disks. At each stage of opening, the finer details are successively eliminated, and the volume of the input image is reduced. Successive openings create a decreasing sequence of images, i.e. $f \circ g_1 \supset f \circ g_2 \dots \supset f \circ g_N$. The image volume remaining after each opening constitutes a decreasing sequence which eventually reaches zero, i.e. $\Omega(1) \geq \Omega(2) \geq \dots \geq \Omega(N)$, where $\Omega(j)$ is the image volume left after the j^{th} opening. This sequence is called the size distribution [5].

The normalised size distribution represents the cumulative proportion of the image volume removed after each opening. It is found by dividing the removed area by the original image volume $\Omega(0)$, i.e. $\Phi(n) = 1 - \Omega(n)/\Omega(0)$. This rises monotonically from 0 to 1 as the size of the SE increases, giving a cumulative distribution function (cdf). Its derivative $\Phi'(n) = d\Phi(n)/dn$ is a probability density function (pdf). This pdf is referred to as the *pattern spectrum* in [11]. Since this is a pdf, it possesses statistical moments. These moments can be used for texture classification and analysis. Granulometric moments from the pattern spectrum of the image foreground provide information on object shape, while those from the pattern spectrum of the image background provide information about spatial distribution of the objects. Here we use these moments to predict the time state or class of an image from a sequence of evolving or ordered texture images.

2.2 Grey level co-occurrence matrix (GLCM)

Entry $C(i, j)$ of a GLCM is defined by first specifying a displacement d and angle ϕ , and counting all pairs of pixels separated by distance d and lying on a line at angle ϕ to the reference direction of the image, which have grey levels i and j respectively. The image is often quantised, e.g. to level 8 or 64, before computing the GLCM, to avoid sparsity of the GLCM. The normalised GLCM $p(i, j)$ can be obtained by dividing $C(i, j)$ by the sum of its entries, as

$$p(i, j) = C(i, j) / \sum_{k=1}^P \sum_{l=1}^P C(k, l),$$

where P is the number of grey levels. GCLMs capture properties of a texture but are not directly useful for further analysis, such as comparison of two textures. Various texture features may be computed from the GLCM for more compact texture representation [9, 16, 4], including:

1. Maximum probability : $\max_{(i,j)} p(i, j)$
2. Energy : $\sum_{i=1}^P \sum_{j=1}^P p(i, j)^2$
3. Entropy : $-\sum_{i=1}^P \sum_{j=1}^P p(i, j) \log p(i, j)$
4. Contrast: $\frac{1}{(P-1)^2} \sum_{i=1}^P \sum_{j=1}^P (i-j)^2 p(i, j)$
5. Homogeneity: $\sum_{i=1}^P \sum_{j=1}^P \frac{p(i,j)}{1+|i-j|}$
6. Correlation : $\frac{1}{\sigma_i \sigma_j} \sum_{i=1}^P \sum_{j=1}^P (i - \mu_i)(j - \mu_j) p(i, j)$, where $\mu_i = \sum_{j=1}^P i \sum_{j=1}^P p(i, j)$, $\mu_j = \sum_{i=1}^P j \sum_{j=1}^P p(i, j)$, $\sigma_i = \sum_{i=1}^P (i - \mu_i)^2 \sum_{j=1}^P p(i, j)$, and $\sigma_j = \sum_{j=1}^P (j - \mu_j)^2 \sum_{i=1}^P p(i, j)$.

2.3 Classification methods

Either granulometric moments or GLCM features are modelled as a function of time using a cubic polynomial regression approach. Let $Y_i(t)$, $i = 1, 2, \dots, p$, and $t = 1, 2, \dots, T$ be the average of the i^{th} feature for the t^{th} time or class (averaged over the training feature set). The cubic polynomial regression can be written as:

$$Y_i(t) = \beta_0^{(i)} + \beta_1^{(i)} * t + \beta_2^{(i)} * t^2 + \beta_3^{(i)} * t^3 + \xi_i, \quad (1)$$

where the β terms are estimated using least squares and the ξ_i are error terms. We have built one such model for each feature used and combined them together. For a single feature i the model will be of the form:

$$\begin{bmatrix} Y_i(1) \\ Y_i(2) \\ \vdots \\ Y_i(T) \end{bmatrix} = \begin{bmatrix} \beta_0^{(i)} + \beta_1^{(i)} t_1 + \beta_2^{(i)} t_1^2 + \beta_3^{(i)} t_1^3 \\ \beta_0^{(i)} + \beta_1^{(i)} t_2 + \beta_2^{(i)} t_2^2 + \beta_3^{(i)} t_2^3 \\ \vdots \\ \beta_0^{(i)} + \beta_1^{(i)} t_T + \beta_2^{(i)} t_T^2 + \beta_3^{(i)} t_T^3 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_T \end{bmatrix}.$$

For p such features, the combined fitted model relating each feature to time or class is:

$$\begin{bmatrix} \hat{Y}_1(t) \\ \hat{Y}_2(t) \\ \vdots \\ \hat{Y}_p(t) \end{bmatrix} = \begin{bmatrix} \hat{\beta}_0^{(1)} + \hat{\beta}_1^{(1)} t + \hat{\beta}_2^{(1)} t^2 + \hat{\beta}_3^{(1)} t^3 \\ \hat{\beta}_0^{(2)} + \hat{\beta}_1^{(2)} t + \hat{\beta}_2^{(2)} t^2 + \hat{\beta}_3^{(2)} t^3 \\ \vdots \\ \hat{\beta}_0^{(p)} + \hat{\beta}_1^{(p)} t + \hat{\beta}_2^{(p)} t^2 + \hat{\beta}_3^{(p)} t^3 \end{bmatrix}, \quad (2)$$

or equally

$$\begin{bmatrix} \hat{Y}_1(t) \\ \hat{Y}_2(t) \\ \vdots \\ \hat{Y}_p(t) \end{bmatrix} = \begin{bmatrix} \hat{\beta}_0^{(1)} \\ \hat{\beta}_0^{(2)} \\ \vdots \\ \hat{\beta}_0^{(p)} \end{bmatrix} + \begin{bmatrix} \hat{\beta}_1^{(1)} & \hat{\beta}_2^{(1)} & \hat{\beta}_3^{(1)} \\ \hat{\beta}_1^{(2)} & \hat{\beta}_2^{(2)} & \hat{\beta}_3^{(2)} \\ \vdots & \vdots & \vdots \\ \hat{\beta}_1^{(p)} & \hat{\beta}_2^{(p)} & \hat{\beta}_3^{(p)} \end{bmatrix} \begin{bmatrix} t \\ t^2 \\ t^3 \end{bmatrix}.$$

Using matrix-vector notation this can be written as:

$$\begin{bmatrix} \hat{\mathbf{Y}} - \hat{\beta}_0 \end{bmatrix} = \begin{bmatrix} \hat{\beta}_1 & \hat{\beta}_2 & \hat{\beta}_3 \end{bmatrix} \begin{bmatrix} t \\ t^2 \\ t^3 \end{bmatrix} \quad (3)$$

or $\begin{bmatrix} \hat{\mathbf{Y}} - \hat{\beta}_0 \end{bmatrix} = \hat{\mathbf{B}}\mathbf{T}$.

Pre-multiplying by $(\hat{\mathbf{Y}} - \hat{\beta}_0)^T$, we get

$$(\hat{\mathbf{Y}} - \hat{\beta}_0)_{(1 \times p)}^T (\hat{\mathbf{Y}} - \hat{\beta}_0)_{(p \times 1)} = (\hat{\mathbf{Y}} - \hat{\beta}_0)_{(1 \times p)}^T \hat{\mathbf{B}}_{(p \times 3)} \mathbf{T}_{(3 \times 1)}. \quad (4)$$

Equation (4) can be written as $at^3 + bt^2 + ct + d = 0$. A positive real root of this equation is used as the predicted time or class. Where there is more than one positive real root we choose the smallest one (as was appropriate in all our training examples). If none of the roots are positive and real the method fails to predict time, and we choose the first time point or class as the prediction.

We compare performance of this regression approach with three other classifiers used to classify objects into mutually exclusive and exhaustive classes based on a set of measurable object features. Linear discriminant analysis (LDA) [8] is a statistical technique which assigns a new feature vector x to the class which has highest posterior probability, assuming the class conditional distributions are multivariate normal with common covariance matrix. The proportions of the data in each time or class are used as prior probabilities. A feed-forward single hidden layer neural network [10] is also used, with 5 units in the hidden layer chosen for optimal performance (in training we tried between 2 and 7 units). Since the performance of SVMs is greatly affected by the choice of kernel function and its associated parameter values, different kernels with different combinations of parameter values and the cost constraint were experimented with to tune the SVM for better performance. One-to-one classification is used here as a means of multi-class classification. A detailed account of SVMs can be found in [10].

The prediction abilities of the classifiers are assessed using proportion of misclassifications and mean absolute error (MAE), defined as:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |t_{pred}^i - t_{act}^i|, \quad (5)$$

where n is the number of images for which time or class is to be predicted, and t_{pred}^i and t_{act}^i are respectively the (rounded) predicted and actual state of time or class of image i .

3. APPLICATION TO REAL IMAGES

3.1 Corrosion images

We applied our methodology to a set of real corrosion images generated in [15]. These are images of a steel plate corroded over a period of 10 consecutive days and photographed daily. The original colour images are of size 1400^2 and were converted to grey scale. To obtain training and test sets, 10 non-overlapping images of size 256^2 were extracted from each of the grey scale images at times $t = 1, \dots, 10$. So there are a total of 100 sample images, 10 for each time point.

As the images were converted from colour images showing substantial intensity variations, image pre-processing

was needed. Spots within the images are darker than the background, so the bottom-hat transform is appropriate for pre-processing. Subtracting the original image from the opened image produces the bottom-hat transform

$$f \hat{\bullet} g = (f \bullet g) - f.$$

We use a disk SE of increasing size as g , since the corrosion spots increase in size over time. A wide range of radius values were tested and the most appropriate ones (6, 7, 8, ..., 15) were chosen for times 1 through 10 to preserve best the sizes and shapes of the textures. Figure 1 shows some of the extracted grey scale corrosion images and their bottom-hat transformed images.

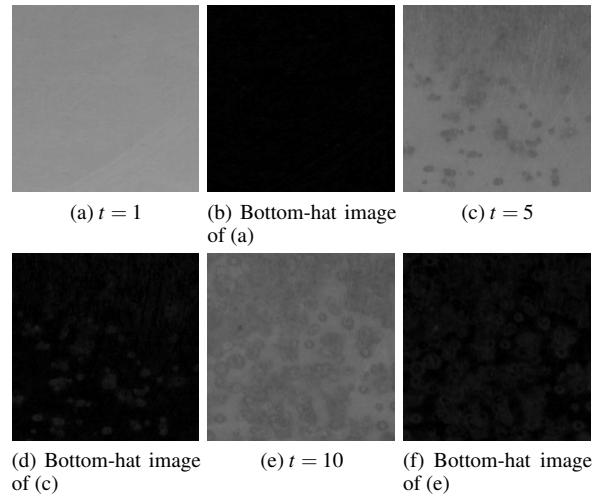


Figure 1: Grey scale corrosion sub-images of size 256^2 and their bottom-hat transformed images at times $t = 1$, $t = 5$ and $t = 10$.

The first four granulometric moments computed from the bottom-hat transformed images using square and disk SEs are shown in Table 1. Average granulometric mean and standard deviation (sd) using both SEs clearly increase over time, while skewness decreases. Kurtosis does not vary much over time, however in the results below we use all 4 moments from both SEs.

Since there are 10 sub-images at each time point $t = 1, 2, \dots, 10$, the moments data form a 100×8 matrix. A randomly selected sub-set of 70% of the moments are used to fit the cubic polynomial regression in Equations 1–3 and the rest are used to predict as in Equation 4. The procedure was repeated 10 times and average results were obtained. The same approach was used for LDA, the neural network and the SVM as well.

3.2 Tea images

Each original image is a colour image of size 2000×3008 . These represent eight different grades of tea [2], i.e. BOPL (Broken Orange Pekoe Large), BOP (Broken Orange Pekoe), BOPSM (Broken Orange Pekoe Small), BP (Broken Pekoe), PF (Pekoe Fannings), PD (Pekoe Dust), OF (Orange Fannings), and Dust, and the approximate diameter in mm of the granules are 2.0, 1.7, 1.3, 1.0, 0.5, 0.355, 0.25 and *Not specific* respectively. We label these as classes 8 to 1, in order of granule size. We obtained training and test sets by extracting

Table 1: First four granulometric moments for the bottom-hat transformed corrosion images using square and disk SEs.

Time	Moments using square SE			
	mean	sd	skewness	kurtosis
1	6.6565	4.5828	0.0032	-2.9977
2	7.5256	5.4877	0.0022	-2.9989
3	8.8921	6.2471	0.0014	-2.9993
4	10.1779	7.2285	0.0009	-2.9996
5	11.2492	7.9117	0.0007	-2.9997
6	12.4619	8.9238	0.0005	-2.9998
7	13.7676	10.1043	0.0004	-2.9999
8	14.8493	11.2805	0.0003	-2.9999
9	17.6761	12.6737	0.0002	-3.0000
10	17.8071	13.8678	0.0002	-3.0000
	Moments using disk SE			
	mean	sd	skewness	kurtosis
1	3.2607	2.6405	0.0227	-2.9817
2	3.7773	3.1449	0.0141	-2.9908
3	4.5955	3.5949	0.0085	-2.9945
4	5.3489	4.1534	0.0054	-2.9969
5	5.9891	4.5508	0.0041	-2.9978
6	6.6898	5.0891	0.0029	-2.9986
7	7.4305	5.6975	0.0020	-2.9991
8	7.9708	6.1374	0.0016	-2.9993
9	9.5363	6.7399	0.0010	-2.9995
10	9.5141	7.4247	0.0008	-2.9997

256² size images from each of the original images and converted them to grey scale. Fifty non-overlapping sub-images were extracted from one image from each class, giving 400 sample images. One sub-image from each class is shown in Figure 2, to show the increasing size of the tea granules over classes.

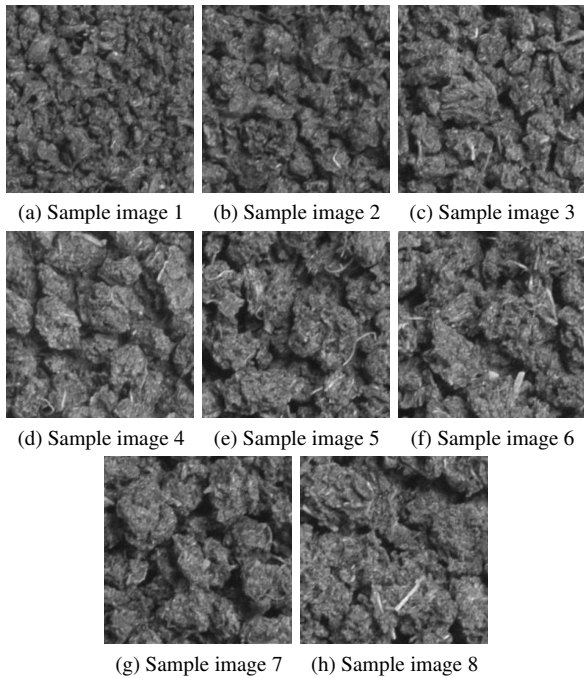


Figure 2: Sample grey scale tea images of size 256², one from each of the eight classes.

There is substantial variation of intensity within the images. To improve this, a top-hat transformation was applied as pre-processing. Since granule size increases over classes, a disk SE of increasing size is used in the top-hat transform (with radii 13, 15, 17, ..., 27). We applied granulometry on the transformed images using square and disk SEs separately and computed the first four PS moments from the pattern spectrum from each SE. Average PS moments are calculated using all sub-images from each class to give an 8 × 8 data matrix. Again we implemented the regression approach, LDA, neural network and SVM classifiers and compared their performance.

4. ANALYSIS AND FINDINGS

For all classifiers 70% of the moments data, randomly selected, was used for training and the rest for testing. This was repeated 10 times and results were averaged.

For the corrosion images, Table 2 shows the performance of all classifiers using the granulometric moments. No image was misclassified by more than one time point, so the MAE and misclassification rate are equal. For the regression approach, 29 images are misclassified out of 300, giving 90.27% correct classification. The SVM using the radial basis kernel is the best, with an average correct classification rate of 99.33% (using cost = 100 and kernel parameter $\gamma = 0.9$). Similar results can be obtained for many parameter settings. LDA is the second best classifier with 93.67% correct classification. The neural network with optimum parameter setting produces a higher misclassification rate. The regression approach is better than the neural network.

Table 2: Time-wise MAE or misclassification rate using all classifiers with 8 granulometric moments, for the corrosion images.

Time	Error rate			
	Regression	SVM	LDA	NNET
1	0.0000	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.1333	0.0333
3	0.0000	0.0000	0.0000	0.0000
4	0.0333	0.0000	0.0000	0.1333
5	0.1000	0.0000	0.1333	0.2000
6	0.0000	0.0000	0.1000	0.3333
7	0.0067	0.0000	0.0667	0.2667
8	0.0333	0.0000	0.1000	0.2667
9	0.4667	0.0000	0.0667	0.3333
10	0.3333	0.0667	0.0333	0.1000
Overall	0.0973	0.0067	0.0633	0.1667

Using the 6 GLCM features at quantisation level 8, the regression approach produces only 15.67% correct classification, SVM produces 88%, LDA produces 22% and the neural network gives 23% correct classification (Table 3). Using the GLCM features some predicted times are more than one unit away from the actual time state. Quantisation level 64 is better for all except the regression approach, which only has 15% correct classification, while LDA has 37.67%, the neural network has 36.33% and SVM has 93.33% correct classification (similar or worse results were obtained using no quantisation). There was no clear advantage in using more than these 6 GLCM features. However, the granulometric moments provide much better results for all classifiers.

Table 3: Error rates for all classifiers using 6 GLCM features at quantisation levels 8 and 64, for the corrosion images.

Time	Quantisation level 8			
	Regression	SVM	LDA	NNET
1	0.1000	0.2333	0.8667	0.9000
2	0.4000	0.2333	0.6000	0.7667
3	0.9333	0.0000	0.9000	0.6667
4	1.0000	0.1667	0.9000	0.9000
5	1.0000	0.1333	0.7667	0.8333
6	1.0000	0.1000	1.0000	0.9333
7	1.0000	0.0667	0.8667	0.8333
8	1.0000	0.2333	0.9000	0.9000
9	1.0000	0.0000	0.5000	0.4667
10	1.0000	0.0333	0.5000	0.5000
Overall	0.8433	0.1200	0.7800	0.7700
	Quantisation level 64			
	REG	SVM	LDA	NNET
1	0.9333	0.1000	0.2333	0.2000
2	0.7333	0.1333	0.9000	0.8000
3	0.8000	0.0000	0.6667	0.8333
4	0.7333	0.0667	0.8333	0.9000
5	0.9333	0.0333	0.8000	0.9333
6	0.5333	0.2333	0.8000	0.8333
7	0.9000	0.0333	0.6000	0.5000
8	0.9667	0.0667	0.3667	0.5000
9	0.9667	0.0000	0.3667	0.3333
10	1.0000	0.0000	0.6667	0.5333
Overall	0.8500	0.0667	0.6233	0.6367

The new regression-based method for classifying images of corrosion to a point in time substantially improves on the results from the approach in [14, 15, 7], where the lowest misclassification rate for the corrosion images was as high as 48%.

For the tea images, the regression classifier produced 89.50% correct classification using granulometric moments, however 100% correct classification can be obtained using all other classifiers if appropriate parameter values are chosen. All the classifiers outperform previous results for these images as reported in [2]. Our highest misclassification rate of 10.50% is for the regression approach, using top-hat transformed images, whereas the lowest misclassification rate in [2] was 20%.

5. SUMMARY AND CONCLUSIONS

In this paper, a regression-based classification approach was applied to two sets of real images and improved results are obtained compared to the existing published work on these images. Increasing the radius of the disk SE in the bottom- or top-hat transform of the images is of crucial importance, as granulometric features computed from the hat transformed images obtained using the same size disk SE over all time points or classes produced very high classification error for all classifiers. Our results are not exactly comparable to [14] and [2], as we have extracted our own sub-images for algorithm development and testing. Nonetheless we conclude that extracting shape-based information from the images directly by use of morphological techniques provides very useful features compared to GLCM features for texture classification in any of a range of classifiers.

REFERENCES

- [1] Batman, S. and Dougherty, E.R. (1997). Size distributions for multivariate morphological granulometries: texture classification and statistical properties. *Optical Engineering*, 36(5), 1518-1529.
- [2] Borah, S., Hines, E.L. and Bhuyan, M. (2007). Wavelet transform based image texture analysis for size estimation applied to the sorting of tea granules. *Journal of Food Engineering*, 79, 629-639.
- [3] Chanda, B. and Majumder, D. (1988). A note on the use of the gray level co-occurrence matrix in threshold selection. *Journal of Signal Processing*, 15(2), 149-167.
- [4] Clausi, D.A. (2002). An analysis of co-occurrence texture statistics as a function of grey level quantization. *Canadian Journal of Remote Sensing*, 28(1), 45-62.
- [5] Dougherty, E.R. and Lotufo, R.A. (2003). *Hands-on Morphological Image Processing*. SPIE Press, Washington, USA.
- [6] Gonzalez, R.C. and Woods, R.E. (2008). *Digital Image Processing, Third Edition*. Prentice Hall, New Jersey.
- [7] Gray, A.J., Marshall, S. and McKenzie, J. (2006). Modeling of evolving textures using granulometries. In Marshall, S. and Sicuranza, G.L. (eds.), *Advances in Non-linear Signal and Image Processing*, EURASIP Series on Signal Processing and Communications, Hindawi Publishing Corporation, New York.
- [8] Hand, D.J. (1981). *Discrimination and Classification*. John Wiley and Sons Ltd, Chichester, UK.
- [9] Haralick, R.M., Shanmugan, K. and Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(6), 610-621.
- [10] Izenman, A.J. (2008). *Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning*. Springer, USA.
- [11] Maragos, P. (1989). Pattern spectrum and multiscale shape representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), 701-716.
- [12] Masotti, M. and Campanini, P. (2008). Texture classification using invariant ranklet features. *Pattern Recognition Letters*, 29(14), 1980-1986.
- [13] Matheron, G. (1975). *Random Sets and Integral Geometry*. John Wiley and Sons Ltd, New York.
- [14] McKenzie, J., Marshall, S., Gray, A.J. and Dougherty, E.R. (2003). Morphological texture analysis using the texture evolution function. *International Journal of Pattern Recognition and Artificial Intelligence*, special issue on Quantitative Image Morphology, 17(2), 167-185.
- [15] McKenzie, J. (2004). *Classification of dynamically evolving textures using evolution functions*. Ph.D. Thesis, Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow.
- [16] Soh, L-K. and Tsatsoulis, C. (1999). Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. *IEEE Transactions on Geoscience and Remote Sensing*, 37(2), 780-795.