

New Distance Measures Applied to Marble Classification

J. Caldas Pinto, J.M. Sousa, and H. Alexandre

Technical University of Lisbon, Instituto Superior Técnico
Dept. of Mechanical Engineering, GCAR/IDMEC
1049-001 Lisbon, Portugal
{jcpinto, j.sousa}@dem.ist.utl.pt

Abstract. Automatic marble classification based on their visual appearance is an important industrial issue, but due to the presence of randomly distributed high number of different colors and its subjective evaluation by human experts, the problem remains unsolved. In this paper, several new measures based on similarity tables built by human experts are introduced. They are used to improve the behavior of some clustering algorithms and to quantitatively characterize the results, increasing the correspondence of the measures to the visual appearance of the results. The obtained results show the effectiveness of the proposed methods.

1 Introduction

Ornamental stones are quantitatively characterized by properties such as geological-petrographical and mineralogical composition, or mechanical strength. However, the properties of such products differ not only in terms of type, but also in terms of origin, and their variability can also be significant within the same deposit or quarry. Though useful, these methods do not fully solve the problem of classifying a product whose end-use makes appearance so critically important. Appearance is conditioned not only by the kind of stone, but it also depends on the subjective evaluation of “beauty”. Traditionally, the selection process is based on human visual inspection, given a subjective characterization of the materials’ appearance, instead of an objective, reliable measurement of the visual properties, such as color, texture, shape and dimensions of their components. Yet, quality control is essential to keep marble industries competitive: shipments of finished products (e.g. slabs) must be of uniform quality, and the price demanded for a product, particularly a new one, must somehow be justified. Thus, it is very important to have a tool for the objective characterization of appearance.

In this paper, we are concerned with marble classification. Several clustering techniques for this purpose are presented and discussed. In order to support these algorithms and to better emulate results obtained by human experts, several new distances and measures are proposed. Hence, we introduce a so-called *Weighted Manhattan* (WM) distance, whose weights are determined by a genetic algorithm, which optimizes a distance between marbles based on a similarity table built by human experts. This distance is used to automatically cluster the marbles. The clustering techniques obtain the clusters based on a set of features. These features are derived by a quadtree based segmentation analysis of marble images, as previously presented and discussed in [6,7]. The clustering results are evaluated by the proposed new error measures, which are motivated by

the fact that standard error measures do not satisfactory agree with the classification results obtained by visual inspection of human experts. The proposed measures are the *similarity table* and the *similarity score*, which are based on expert evaluations.

This paper is organized as follows. In Section 2 the distance measures are presented, and the weighted Manhattan distance, the similarity error and the similarity score, which are all proposed in this paper, are described. Section 3 presents a short description of the clustering algorithms used in this paper, namely, simulated annealing, fuzzy c-means, neural networks, and Takagi-Sugeno fuzzy models optimized by genetic algorithms. The application to real marbles data is discussed in Section 4. This section presents the results demonstrating the validity of the approaches introduced in this paper. Section 5 concludes the paper, giving guidelines for futures developments in this important area.

2 Distance Measures

The evaluation of the obtained results can be a significant problem in situations where the visual appearance dictates the qualitative results. Sometimes results are visually very good but this fact is not revealed using standard error measures. On the other hand, a team of experts can built a so-called *similarity table* where, for each marble, those that could be considered very similar in appearance, from the available collection, are indicated. In the application in this paper, this number varies from 2 up to 10 marbles with an average of 4. The marbles considered to bo similar to a given marble are called *neighbors*. The table containing the marbles that are similar to a given example is called the *complete similarity* (CS) table. From this table another was derived, with a maximum number of 4 neighbors, and where a weight ω reflecting the number of votes received from the experts was added to each neighbor. This table is named *weighted similarity* (WS) table.

2.1 Weighted Manhattan Distance

The weighted similarity table can be used to develop a new distance measure, the so-called *Weighted Manhattan* distance, which can find correct neighbors for new marbles in a simple way. The WM distance between the points x_m and x_n is given by:

$$d_w = \sum_{i=1}^N \alpha_i |x_{m_i} - x_{n_i}|^{\beta_i}, \tag{1}$$

where $i = 1, \dots, N$ are the coordinates of the space (features) with dimension N . The weights α_i and β_i are determined by optimizing the cost function

$$J^{(k)} = \sum_{i=1}^{N_d} \sum_{v=1}^{n_{\max}} \omega_{iv} S(n_{iv} - n_{iv}^{(k)}) \tag{2}$$

where k is the iteration of an optimization algorithm, N_d is the total number of data points, n_{\max} is the maximum number of neighbors defined for each data point, n_{iv} is the v neighbor of data point i , $n_{iv}^{(k)}$ is the v neighbor defined at step k , and ω_{iv} is the

weight, defined by the experts, associated to the v neighbor of data point i . The function S is simply given by

$$S(n_{iv} - n_{iv}^{(k)}) = \begin{cases} 0, & \text{if } n_{iv} \neq n_{iv}^{(k)} \\ 1, & \text{otherwise} \end{cases} \quad (3)$$

Note that in this paper the data points are marbles, N_d is equal to 112, and n_{\max} was settled to 4. A standard genetic algorithm based on real code [5] was used to optimize the weights α_i and β_i in (1). Due to lack of space, we do not present the genetic algorithm parameterization in this paper.

2.2 Similarity Error and Similarity Score

This paper intends to cluster a given collection of marbles from a set of features, in order to be able to classify these marbles. Besides the visual appearance of the results, which is in general very good, it is important to evaluate quantitatively the classification results in order to compare different algorithms. However, the most common error measures based on the differences between the original marble classification and the attributed one are not consistent with the visual quality of the clusters. This is mainly due to the difficulty of experts in attributing a certain class to a given marble. Most marbles could be classified in a different, although similar, way. Based on the importance of the visual appearance, the CS table can be used to evaluate the clustering algorithms. Therefore, this paper introduces two new measures, which are called *similarity score* and *similarity error*, respectively.

Let V_{x_i} be the set of neighbors of the marble x_i and C_j be the cluster j with n_j elements. A score for the cluster C_j can be defined as

$$SC_j = \sum_{i=1}^{n_j} \#(V_{x_i} \cap C_j). \quad (4)$$

The global score of a given clustering algorithm for the total number of clusters n_c is given by:

$$SC = \sum_{i=1}^{n_c} SC_j \quad (5)$$

On the other hand, a marble is wrongly integrated in a cluster if there are no marbles of its neighborhood in that cluster. So a *similarity error* measure can be defined as follows:

$$SE_j = \sum_{i=1}^{n_j} f(V_{x_i} \cap C_j), \quad (6)$$

where

$$f(V_{x_i} \cap C_j) = \begin{cases} 1, & \text{if } V_{x_i} \cap C_j = \emptyset \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

And the global error of a given clustering algorithm for the total number of clusters n_c is given by:

$$SE = \sum_{i=1}^{n_c} SE_j \quad (8)$$

This measure can be normalized in order to facilitate the comparisons. The normalization is given by:

$$se = \frac{SE}{N_d} \quad (9)$$

The similarity score SC and the similarity error SE are related but not dual, as it will be clear in the results presented in Section 4.2.

3 Clustering Algorithms

This paper uses both supervised and unsupervised clustering methods. The unsupervised methods do not use the training set classification to derive the proper clusters. This type of algorithms has a cost function that must be optimized. The classification of a given data point (marble) is achieved through the distance of its features to the respective clusters. This paper uses two types of unsupervised clustering algorithms: simulated annealing and fuzzy c-means. In the supervised methods, the classification given in the training set is used to derive the clusters. Thus, the classification is achieved through the simulation of the derived model, where the inputs are the marbles to be classified. The supervised methods applied in this paper are neural networks, and fuzzy models optimized by a genetic algorithm.

Let $\{\mathbf{x}_1, \dots, \mathbf{x}_{N_d}\}$ be a set of N_d data objects where each object is an instance represented by the vector $\mathbf{x}_i \in \mathbb{R}^N$, which is described by a set of N features. The set of data objects can then be represented as a $N_d \times N$ data matrix \mathbf{X} . The clustering algorithms used in this paper determine a partition of \mathbf{X} into C clusters.

3.1 Simulated Annealing

This optimization algorithm simulates the annealing process, by reproducing the crystallization process of particles. These particles move during the solidification process, which occurs when the temperature decrease. In the final stage, the system reaches the minimal energy configuration [8]. The process starts at an initial temperature, which should consider the number of changes that increases the system energy, in order to allow the system to escape from local minima. This algorithm defines the proportion for acceptable changes that increases the system's energy. The cost function is represented by the Euclidean, Manhattan, or weighted Manhattan distances between the data points and the cluster centers. This cost function indicates if the system is converging or not to a lower energy state. The number of iterations to accomplish at each temperature is proportional to the number of elements. This parameter is proportional to the dimension of the data set N_d . The temperature T must decrease between the several energy levels. For consecutive levels, the temperature decrement is given by

$$T^{(k+1)} = d \cdot T^{(k)} \quad (10)$$

The *stop criterion* can be given by a predetermined number of temperature decreases or when after some consecutive temperature decreases the final cost does not change.

3.2 Fuzzy C-Means

Considering the set of data objects \mathbf{X} , the fuzzy c-means algorithm determine a fuzzy partition of \mathbf{X} into C clusters by computing a $N_d \times C$ partition matrix \mathbf{U} and the C -tuple of corresponding cluster prototypes [2]:

$$\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_C]. \quad (11)$$

Most often, the cluster prototypes are points in the cluster space, i.e. $\mathbf{v}_i \in \mathbb{R}^N$. The elements $u_{ik} \in [0, 1]$ of \mathbf{U} represent the membership of data object \mathbf{x}_k in cluster i . This paper uses the standard fuzzy c-means algorithm, as described in [2,1]. Note however, that two distance measures are used; the Euclidean and the WM distances.

3.3 Neural Networks

Neural networks have been largely used as input-output mapping for different applications including modeling, control and classification [4]. The main characteristics of a neural network are its parallel distributed structure and its ability to learn, which produces reasonable outputs for inputs not encountered during training. Moreover, the structure can be chosen to be simple enough to compute the output(s) from the given input(s) in very low computational time.

The neural network used in this paper has hidden-layers with hyperbolic tangent activation functions and a linear output layer. The network must have few neurons in the hidden-layers, and is trained using the resilient backpropagation algorithm [3].

3.4 Fuzzy Models

Fuzzy models have gained in popularity in various fields such as control engineering, decision making, classification and data mining [10]. One of the important advantages of fuzzy models is that they combine numerical accuracy with transparency in the form of rules. Hence, fuzzy models take an intermediate place between numerical and symbolic models. A method that has been extensively used for obtaining fuzzy models is fuzzy clustering. Several fuzzy modeling approaches to classification based on fuzzy clustering, have been compared in [11]. From the ones compared the optimization of Takagi-Sugeno fuzzy models using a genetic algorithm (GA) as described in [9] proved to be the best, and as so, this algorithm is applied in this paper.

4 Application: Marble Classification

4.1 Parameters of the Algorithms

The clustering algorithms are tested for two situations: clustering veins for the color with more elements, where the training set has 29 marbles and the test set 14 marbles, and clustering colors, where the training set has 69 marbles and the test set 43 marbles.

First, simulated annealing and fuzzy c-means was applied using the Euclidian distance and the weighted similarity distance. These classifiers used 6 and 3 clusters to

classify marbles color and marbles veins, respectively. In the SA algorithm, the acceptable change used to compute the initial temperature is 0.3, the constant used to compute the number of iterations to accomplish in each temperature is set to 5, the parameter d in (10) is equal to 0.9; and finally 100 temperature decreases are allowed.

In the feedforward neural networks, the parameters are derived experimentally in order to obtain good classification results. In this paper, three hidden layers with 9, 12 and 9 neurons showed to derive the best classification results. The number of epochs is set to 100, in order to avoid overfitting.

The Takagi–Sugeno fuzzy models returns as output a real number, which must be transformed to a given class c_k . As so, the output of the model y is set to a class as follows:

$$c_k = \text{round}(y). \tag{12}$$

The classification value c_k corresponds to the set $c_k \in \{1, 2, 3\}$ for the case of the veins, and to the set $c_k \in \{1, 2, 3, 4, 5, 6\}$ when the colors are classified.

4.2 Results

Table 1 presents the comparison of the results obtained for vein classification. The error measures are the the following: e – mean classification error, SC – global score of clusters as defined in (5) and se – normalized similarity error, as defined in (9). The clustering algorithms in Table 1 are the simulated annealing (SA), the simulated annealing using the WM distance (SAw), the fuzzy c-means (FC), the fuzzy c-means using the WM distance (FCw), neural networks (NN) and the fuzzy model optimized by genetic algorithms (FGA). Table 2 presents the comparison of the results obtained for color classification. The notation used in Table 1 stands also for Table 2.

Table 1. Comparison of the results obtained for vein clustering.

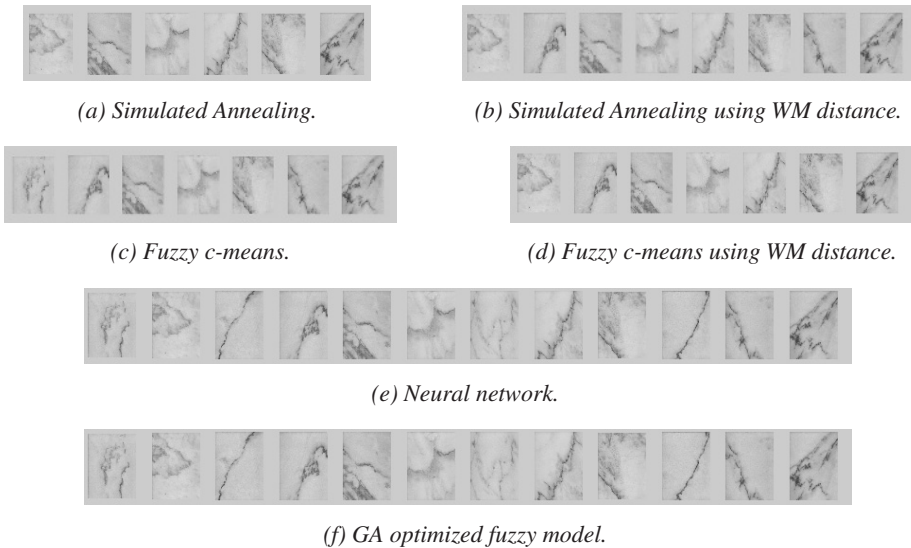
	e		se		SC	
	train data	test data	train data	test data	train data	test data
SA	0.41	0.42	0.00	0.14	90	23
SAw	0.38	0.50	0.03	0.00	112	23
FC	0.41	0.36	0.03	0.29	89	21
FCw	0.38	0.50	0.03	0.00	112	23
NN	0.10	0.21	0.03	0.14	130	22
FGA	0.10	0.50	0.03	0.07	130	19

In both tables it is clear that the results of the unsupervised methods are not so good as the ones using the supervised methods, especially when evaluated with traditional measures. However, the measures se and SC proposed in this paper allow to conclude that in general the WM distance improved the performance of the unsupervised algorithms. Thus, the apparent big differences between supervised and unsupervised algorithms are less evident with the introduced measures SC and se . These results are confirmed by

Table 2. Comparison of the results obtained for color clustering.

	e		se		SC	
	train data	test data	train data	test data	train data	test data
SA	0.48	0.67	0.06	0.14	195	93
SAw	0.46	0.53	0.03	0.16	240	87
FC	0.45	0.63	0.06	0.12	197	69
FCw	0.48	0.60	0.06	0.14	234	92
NN	0.13	0.26	0.09	0.07	254	81
FGA	0.10	0.44	0.01	0.16	268	52

visualizing the marbles. Due to the lack of space, we present only the visual results of one clustered region in the vein clustering presented in Table 2. The marble classification is presented in Figure 1.

**Fig. 1.** Example of marbles veins clusters: results for one region in the training set.

5 Conclusions

This paper deals with the problem of clustering natural surfaces based on their visual appearance (marble classification in this case). Standard classification error measures do not satisfactory correspond to the classification results obtained by visual inspection of

human experts. Therefore, new measures are proposed in this paper, namely a *similarity table* and a *similarity score*, which are both based on expert evaluations. Further, we also introduced the so-called weighted Manhattan distance, which is based on a weighted expert table, and tuned through the use of genetic algorithms. This distance was applied to simulated annealing and fuzzy c-means clustering, and in general it improved the clustering results. Two supervised techniques were also tested, neural networks and a fuzzy model optimized by genetic algorithms. All the tested algorithms have shown that the set of features used to describe the marbles are quite adequate for clustering based on visual appearance.

Future work includes further studies in the weighted Manhattan distance, further evaluation and possible improvements of the introduced new measures, and finally a possible combination of clustering techniques to obtain a reliable marble classification system.

Acknowledgements. This research is partially supported by the “Programa de Financiamento Plurianual de Unidades de I&D (POCTI), do Quadro Comunitário de Apoio III”, and by the FCT project POCTI/2057/1995 - CVAM, 2nd phase, Ministério do Ensino Superior, da Ciência e Tecnologia, Portugal.

References

1. R. Babuška. *Fuzzy Modeling for Control*. Kluwer Academic Publishers, Boston, 1998.
2. J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Function*. Plenum Press, New York, 1981.
3. C. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, New York, 1995.
4. S. Haykin. *Neural Networks: A Comprehensive Foundation*. Prentice-Hall, Upper Saddle River, NJ, second edition, 1999.
5. Z. Michalewicz. *Genetic Algorithms + Data Structures = Evolution Programs*. Springer, Berlin, third, revised and extended edition, 1999.
6. J. C. Pinto, A. Marcolino, and M. Ramalho. Clustering algorithm for colour segmentation. In *Proc. of V Ibero-American Symposium On Pattern Recognition, SIARP'2000*, pages 611–617, Lisbon, Portugal, September 2000.
7. J. C. Pinto and J. Sousa. Comparison of fuzzy clustering and quadtree methods applied to color segmentation. In *Proceedings of 12th Portuguese Conference on Pattern Recognition, RecPad'2002*, pages Session: Pattern Recognition, pp. 1–4, Aveiro, Portugal, June 2002.
8. P. Salamon, P. Sibani, and R. Frost. *Fact, Conjectures, and Improvements for Simulated Annealing*. SIAM, Philadelphia, USA, 2002.
9. M. Setnes and H. Roubos. GA-fuzzy modeling and classification: complexity and performance. *IEEE Transactions on Fuzzy Systems*, 8(5):516–524, October 2000.
10. J. Sousa and U. Kaymak. *Fuzzy Decision Making in Modeling and Control*. World Scientific Pub. Co., Singapore, 2002.
11. J. M. Sousa, U. Kaymak, and S. Madeira. A comparative study of fuzzy target selection methods in direct marketing. In *Proceedings of 2002 IEEE World Congress on Computational Intelligence, WCCI'2002, FUZZ-IEEE'2002*, pages Paper 1251, pp. 1–6, Hawaii, USA, May 2002.