DAMAGE DETECTION ON HISTORICAL BUILDINGS USING UNSUPERVISED CLASSIFICATION TECHNIQUES

C. Crespo [a], J. Armesto [a] *, D. González-Aguilera [b], P. Arias [a]

[a] Dep. Natural Resources Engineering and Environmental Engineering, University of Vigo, Campus Universitario As Lagoas –Marcosende s/n 36200 Vigo Spain
[b] Cartographic and Land Engineering Department. University of Salamanca, Hornos Caleros, 50, 05003 Ávila, Spain

Commission V, WG V/3

KEY WORDS: Terrestrial laser scanner, image classification, historical building, damage assessment, intensity data, granitic rock

ABSTRACT:

This paper describes a methodology for the documentation and analysis of historical buildings. Techniques of digital image processing give the possibility to detect damages, such as moisture or biological changes, on surfaces of monuments. Intensity data obtained from laser scanner equipment and colour information gathered with a digital camera were evaluated for damage identification that affects the materials used in several historical buildings, such as granitic rock. Then, digital image processing was used to identify damages over granitic rock in a nondestructive way; in particular, classification methods were applied. Several unsupervised classification algorithms were analyzed. The output data were classified images showing the different kind of damages that affect the granitic rock. Post-analysis of these data allow obtaining thematic maps with the size and position of damages. The data and first results that were obtained to date are described below.

1. INTRODUCTION

In cultural heritage research, the characterization and classification of the conservation state of the materials that constitute historical buildings are essential to study their damage. There are a lot of damages that can affect building materials. The recognition of these damages and their evolution over time is an important task in the documentation work. A manual mapping of these damages is expensive and time-consuming. Besides, data capture depends either on the accessibility of the building, its emplacement and dimensions, and on the way of inspection and analysis. Therefore, the use of terrestrial laser scanning combined with classification techniques can assist in the task, providing thematic maps of damage affecting the building materials of the monuments.

The geometrical information provided by the terrestrial laser scanning has been successfully applied in a number of diverse fields such as architectural heritage, civil engineering, geology or documentation of heritages sites (Armesto et al. 2009, Sharaf et al. 2009). However, the radiometric information, provided by the measurement of the strength of the backscattered signal, is less frequently used. Physical factors such as object reflectivity, which is wavelength dependent; geometric factors including the angle of incidence, the surface roughness and the distance between device and the object and the atmospheric conditions influence in the strength of this signal. In general, this signal (so-called *intensity*) is a qualitative measure of the relative strength of the power of the received signal. Nowadays, one can find works related to the applications of the intensity data (Franceschi et al. 2009, Lichti, 2005).

The analysis of intensity and colour information facilitates the visual interpretation of the images and combination of the image channels. Then, different operations, such as ratio or sum, can be performed between the channels helping to produce a better characterization of the damages to be studied (Lerma, 2001).

The overall objective of image classification procedure is to automatically categorize all pixels in an image into themes. In statistical terms, this means reducing the scale of measurement of a continuous variable (digital level) a categorical scale (in this study, types of damage). Different damage types manifest different combinations of digital levels based on their inherent spectral reflectance and emittance properties. The classification procedure utilizes the pixel-by-pixel spectral information as the basis to generate the clusters for the detection of damages on historical buildings (Lerma 2005). Then, these techniques offer the possibility of detecting damages on building surfaces. Image regions can be assigned to different damages, and the dimension of the damaged region can be quantitatively measured. Therefore, the results of these methods can serve as a basis for damage mapping, delivering quantitative results of the damaged area together with information about the nature of the damage. In the literature, there is available information about the different strategies of classification (Tso, 2001), which vary in complexity. The selection of the most suitable technique depends on the type of images and applications.

In Galicia, an autonomous community in Spain located in the northwest of the Iberian Peninsula, the main building material on historical buildings is granite. Environmental conditions, such as temperature, presence of salt or moisture, contact with

---

* Corresponding author: Tel.: +34 986 813423;
E-mail address: julia@uvigo.es (J. Armesto).

other construction materials and pollution, have caused deterioration of the granite over time. Water promotes biological colonization by mosses, lichens, fungi or moulds (Lisci et al. 2003). Water also increases the ability of air contaminants to produce a hard and blackned layer on the surface of the rock (Marszalek, 2004; Moropoulou, 2003).

The main objective of this study was to use terrestrial laser scanning intensity data and colour information combined with image classification techniques for detecting damage affecting the granite employed on historical buildings.

2. STUDY SITES AND FEATURES

The study site was the ruins of Santo Domingo, located in the city of Pontevedra (Galicia, Spain). The material that constituted this building was granite. This building is located in the city centre. It is exposed to a large number of air contaminants. This exposure along with the particular climatic conditions of this area especially high moisture levels, are responsible for the damage that affect the granitic rock of this building. An expert with previous knowledge of the different types of damage affecting the granitic rock of the ruins of Santo Domingo helped us to select the most interesting areas.

This paper shows the earlier results of the proposed methodology. Then, an area of 3x3 m^2 inside one of the most damaged chapels in the building was chosen as a test area to perform the suggested methodology. This area was chosen because of the wide knowledge of the expert about the damages this area is affected of, which is necessary to accomplish a detailed documentation study. The area has both dry and wet regions. The observation of the area allowed us to identify the following features: (1) areas with high moisture content and (2) remains of lime and mortar.

Figure (1) shows an image of part of the chosen chapel, taken with a digital camera; the analysed region is shown inside the red box.

Figure 1. (left) Image of the damaged chosen chapel, taken with a digital camera; (right) 2D image of the selected region obtained from the 3D information of the colored point cloud.

3. MATERIAL AND METHODS

3.1 Equipment

In this work the terrestrial laser scanner system used was RIEGL Z390i (RIEGL, 2010). This scanner is based in the principle of time of flight and has a laser source that emits pulses with a wavelength of 1,500 nm (infrared). The point measurement range is 11,000 points per second. The signal intensity received by the sensor system is recorded in 8 bits [0 255]. A Nikon D200 camera was mounted on the laser. The software Riscan Pro was used for the acquisition and the processing of the data.

3.2 Data acquisition

Scanning positions were selected according to its technical specifications and by searching the front and the proximity to the studied area in order to minimize the influence of the geometry in data acquisition. The colour information of the area is obtained with the digital camera. This setting was used to perform high resolution inspections of several specific areas in the Ruins of Santo Domingo. Figure (2) shows the position of the laser equipment and the point cloud of the chosen area which will be analysed in this work.
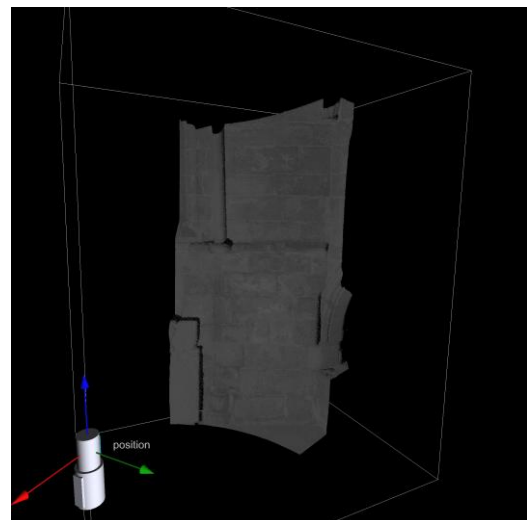

Figure 2. Point cloud represented as a function of the intensity of the reflected laser beam. The number of points of this cloud is 6,737,130 and the angular resolution was 0.020º.

3.3 Data processing

We performed a manual cleaning, which was done on the basis of visual navigation on the point cloud, identifying and eliminating those points that were not part of the object of study. Then, we added colour to the point cloud. Finally, we generated 2D images from the 3D information provided by the point cloud. These 2D images were composed of pixels, each of them defined by its digital number (DN). By this way, we obtained one 2D image displaying the intensity (infrared information) and another displaying the colour information of the features of the area of study.

3.4 Data post-processing: Classification and analysis

Images derived from the combination of the intensity data (infrared) with the red channel of the colour information were used to identify damages on the investigated building surface in an easy and effective way.

For each of the images an unsupervised image classification was performed. This method automatically divides the range of values of the image into clusters according to a statistical procedure, without any predefined knowledge. This strategy assumes that the digital levels of the image are a number of groups (called clusters) more or less distinct as the case. The algorithms used were Isodata algorithm (Tou, 1974), k-means algorithm (Tou, 1974) and Fuzzy k-means algorithm (Bezdek, 1973).

For each classification algorithm, there were 10 predefined classes. After each classification, the non-classified pixels were assigned to the null class and the spectral clusters were merged into meaningful thematic classes. The reference image chosen was a 2D image with colour information in every pixel. To quantify the classification accuracy a number of samples were generated on this reference image selected. A test image was created using a manual assignment between the reference class and the samples. Then, accuracy assessment was conducted over the entire study area by comparing the classified images with the area-truth image. Null class was not taken into account for the accuracy assessment; only the assigned classes after their merge were evaluated. An accuracy report was produced for each case.

To compare the different results of the classification, the overall accuracy was calculated. This statistical measure was calculated from the error matrix from this expression,

$$overall\ accuracy = \frac{\sum_{i=1,n} x_{ii}}{\sum_{i=1,n} \sum_{j=1,n} x_{ij}} \qquad (1)$$

where x_{ii} = diagonal elements of the matrix
x_{ji} = total sampled points

Additionally, in order to assess these results, each of the clusters considered was analyzed through the parameters associated with each class, such as the producer's accuracy (it represents the probability that a randomly chosen pixel belonging to a class has of being correctly assigned to this class) and the user's accuracy (it represents the probability that a randomly chosen pixel classified into a class in the thematic map has of being assigned to this class). These parameters provided analytical results about the comparative performance of classification using different sets of information: intensity or intensity and colour combined information.

4. RESULTS

This section summarizes the results obtained in the post analysis. In this process, the number of classes decreased from 10 classes (the initially defined number of clusters) to 4 classes (including the null class), which were those that had a real meaning (thematic classes). These classes were: lime/mortar remains and ashlars with high moisture level.

Table 1 summarizes the values of overall accuracy (in percentage) obtained for each of the classification algorithms considered and for each data set evaluated. The best results of accuracy were achieved for the Fuzzy k-means algorithm in every image. Adding red information to the intensity information improves the results of the classification.

Overall accuracy (%)			
	Isodata	k-means	Fuzzy k-means
Intensity (IR)	62.14	74.21	78.95
Sum (IR+R)	80.92	84.37	86.25
Ratio (IR/R)	57.48	54.97	58.36

Table 1. Unsupervised classification accuracies achieved for each algorithm considered for every data set.

Figure (3) shows images of intensity (infrared) and the results of the unsupervised classification, after post-analysis, for each algorithm considered in this study. In a visual assessment, the patterns found with k-means and Fuzzy k-means algorithms were different from the pattern found with the Isodata algorithm. The patterns obtained with the first two algorithms had distinct forms representing the areas where remains of lime/mortar appeared accumulated on the ashlars, and the pixels corresponding to these areas were correctly grouped into a class (blue, in figure 3).
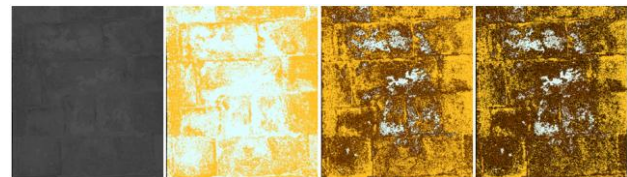

Figure 3. Figure shows (from left to right): Image with infrared information; result Isodata classification; result k-means classification and result Fuzzy k-means classification.

In a quantitative analysis, the best results were obtained for the Fuzzy k-means with an overall accuracy of 78.95% and the worst results belong to the Isodata with an overall accuracy of 62.14%. It was also found that the producer's accuracy for the class remains lime/mortar was 75% and the user's accuracy was 69%, for the Fuzzy k-means algorithm. In the case of class ashlars affected by moisture, values found for the producer's accuracy and the user's accuracy were 70% and 60%, respectively. These results show good accuracies for discriminating ashlars with remains lime/mortar and other ashlars affected by high moisture content. Results for k-means algorithm are similar. Results for Isodata are worse with a 65% and 57% for producer's and user's accuracy, respectively, for the class of lime/remains, and a 60% and 45% for producer's and user's accuracy, respectively, for the class of ashlars affected by moisture. The results of this post-analysis show that the k-means and Fuzzy k-means algorithms distinguish remains lime/mortar from other damages.

Figure (4) shows the image of the sum pixel by pixel of intensity (infrared) and red information and their results in the unsupervised classification, after post-analysis, for each

algorithm considered in this work. In this case, the patterns obtained with the three algorithms are similar. A visual assessment, made by an expert, confirms that the areas with accumulated lime remains on the ashlars are correctly identified in these maps (blue, in figure 4).
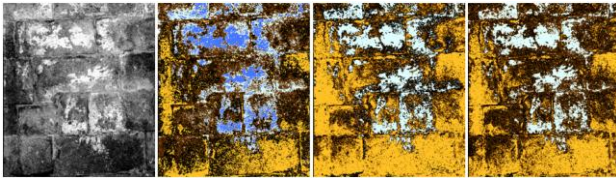

Figure 4. Figure shows (from left to right): Image of the sum of the infrared and red information; result Isodata classification; result k-means classification and result Fuzzy k-means classification.

In a quantitative analysis, the results show high levels of overall accuracy, with values ranging from 80.92% achieved for Isodata algorithm and 86.25% achieved for the Fuzzy k-means algorithm. It was found that the producer's and user's accuracy are high and significant differences were not found among the algorithms used in this study. For Fuzzy k-means, the value for the producer´s accuracy was 95% and for the user's accuracy was 92% for the class remains lime/mortars. These values, for the class of ashlars affected by high moisture content, are worse, with 80% and 72% for producer's and user's accuracy, respectively for the Fuzzy k-means.

These results show that adding the red information to infrared information improves the results of the classification in both quantitative and qualitative results.

Figure (5) shows the image of the ratio pixel by pixel of intensity (infrared) and red information and the results of the unsupervised classification, after post-analysis, for each algorithm considered in this study. In a visual assessment, the pattern found is similar in the three algorithms. In this case, there are pixels that do not belong to the class of remains of lime and mortar (blue in figure 5), however they are included in it, which makes this class extended to ashlars where there was really no such damage.
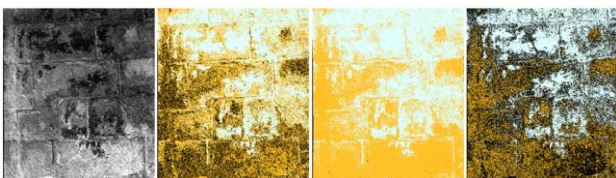

Figure 5. Figure shows (from left to right): Image of the ratio of the infrared and red information; result Isodata classification; result k-means classification and result Fuzzy k-means classification.

In a quantitative analysis, the results show a low value of the overall accuracy, for each algorithm used. In this case, the overall accuracy is bellow 60% for all algorithms (see table 1). The values for the producer's and user's accuracy, for every thematic class considered were below 50%. Then, for producer's accuracy this indicates that less than 5 out of 10 pixels that were in fact affected by the type of damage were included in this class. For user's accuracy, this means that 2 out of 5 pixels classified in a class actually had this type of damage.

5. CONCLUSIONS

The methodology presented includes the following steps, once the images have been obtained: Run unsupervised classification and preview the results. Perform the analysis of the results of the classification, and finally obtain the final report with the statistical analysis.

Unsupervised classification algorithms were chosen because they allow obtaining a classified image automatically and without prior knowledge of the area.

Our results show that the algorithm with the best results is the Fuzzy k-means algorithm, and that the addition of the colour information to the intensity information improves qualitative and quantitative results of the classification.

Our goal has been to begin the development of a working strategy that allows detecting damages on the granite rocks used in construction of historical buildings by means of combination of intensity and colour information and classification techniques.

These first results showed the different response of the ashlars affected by damage to different moisture contents. Therefore, the proposed methodology is a good procedure to diagnosis the damage that affects the buildings materials on historical buildings.

6. ACKNOWLEDGEMENTS

7. REFERENCES

References from journals:

Armesto, J., Ordóñez, C., Alejano, L., Arias, P. 2009. Terrestrial laser scanning used to determine the geometry of a granite boulder for stability analysis purposes. *Geomorphology*, 106, pp. 271-277.

Franceschi, M., Teza, G., Preto, N., Pesci, A., Girardi, S. 2009. Discrimination between marls and limestone using intrensity data from terrestrial laser scanner. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64, pp. 522-528.

Lerma, JL. 2001. Multiband versus multispectral supervised classification of architectural images. *Photogrammetric Record*, 17, pp. 89-101.

Lerma, JL. 2005. Automatic plotting of architectural facades with multispectral images. *Journal of Surveying Engineering-ASCE*, 131, pp. 73-77.

Lichti DD. 2005. Spectral filtering and classification of terrestrial laser scanner point clouds. *Photogrammetric Record*, 20, pp. 218-240.

Lisci, M., Monte, M., Pacini, E. 2003. Lichens and higher plants stone: A review. *International Biodeterioration and Biodegradation*, 51, pp. 1-17.

Moropoulou, A., Polikreti, K., Ruf, V., Deodatis G. 2003. Characterization of building materials, damage assessment and conservation considerations. *Journal of Cultural Heritage*, 4, pp. 101-108.

Sharaf Al-Kheder, Yahya Al-Shawabkeh, Norbert Haala. 2009. Developing a documentation system for desert palaces in Jordan using 3D laser scanning and digital photogrammetry. *Journal of Archaeological Science*, 36, pp. 537-546.

References from Books:

Bezdek JC. 1973. *Fuzzy mathematics in pattern classification*. PhD dissertation Cornell Univ Itheca

Marszalek, M. 2004. *Deterioration of stone in some monuments exposed to air pollution: a Cracow case study*. Air pollution and Cultural Heritage, Taylor and Francis, London. pp. 151-154.

Tou, J., González, R. 1974. *Pattern recognition principles*. Addison-Wesley Publishing Co.

Tso, B., Mather., PM. 2001. *Classification methods for remotely sensed data.* Taylor and Francis, London.

References from websites:

RIEGL Laser Measurements and Systems. High Accuracy, high Resolution 3D Terrestrial Laser Scanner System Z-390i. http://www.riegl.com/uploads/tx_pxpriegldownload (accessed January 2010)