Image-Based Delineation of Built Heritage Masonry for Automatic Classification

Noelia Oses[1] and Fadi Dornaika[2,3]

[1] Fundación Zain Fundazioa, Vitoria-Gasteiz, Spain
[2] University of the Basque Country UPV/EHU, San Sebastian, Spain
[3] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

Abstract. The development of new built heritage assessment protocols that objectivise and standardise the protection process has made it possible to start the work on the algorithms necessary to implement a built heritage analysis and classification ICT tool. The built heritage that will be assessed using these protocols consists of stone masonry constructions. Much of the assessment is carried out through visual inspection. Thus, this process will be automated by applying image processing on digital images of the elements under inspection. Many of the features analysed can be characterised geometrically and are often related to the arrangement of the construction blocks. This paper presents the ground work carried out to make this tool possible: the semi-automatic delineation of the masonry. The validity of this delineation will be shown using the classification results for the analysis of one of the elements assessed in the protocol for masonry bridges.

Keywords: semi-automatic masonry delineation, image processing, classification, built heritage analysis, Hough Transform.

1 Introduction

In the last decade, machine vision techniques have been more and more used in order to assist the whole process of cultural heritage documentation, preservation, and restoration [1–3]. 3D digital modeling for cultural heritage has recently received a lot of attention from the scientific community [4–9]. Many works have been proposed in order to automate the process of 3D modeling of cultural heritage sites or buildings. However, the semantic categorization of the related scenes (images or 3D models) has received little attention. 3D modeling of cultural heritage is useful for visualization, digital archiving and sharing for education, research, and conservation. As can be seen, machine vision and photogrammetry techniques become more and more appealing in order to achieve the related objectives. In the field of geometrical documentation of heritage, research has been carried out in order to automate the processing of 3D point clouds. Rodríguez et al. [10] present a semi-automatic method to draw straight lines in the 3D point cloud. This method requires human interaction to specify the segments of interest in the 2D sections.

M. Kamel and A. Campilho (Eds.): ICIAR 2013, LNCS 7950, pp. 782–789, 2013.

Fundación Zain is developing new built heritage assessment protocols that standardise the protection process and working on a built heritage analysis and classification ICT tool. The objective of this tool is to speed up the protection process. Much of the assessment is carried out through visual inspection. Thus, this process will be automated by applying image processing on digital images of the elements under inspection. Many of the features analysed can be characterised geometrically and are often related to the arrangement of the stone masonry blocks. This paper presents the ground work carried out to make this tool possible: the semi-automatic delineation of the masonry. The validity of this delineation will be shown using the classification results for the analysis of one of the elements assessed in the protocol for masonry bridges. The work on the semi-automatic delineation of masonry presented in this paper is closely related to that of Rodríguez et al. in such way that it could help automate the selection of segments of interest in the 2D sections.

Section 2 describes the problem statement. Section 3 describes the processing steps introduced in order to extract useful information from built heritage images. Section 4 describes some classification results for which the feature extraction was based on the semi-automatic delineation proposed in the previous section. Finally, Section 5 provides the conclusions and future work.

2 Problem Statement: Automatic Delineation and Classification of Masonry

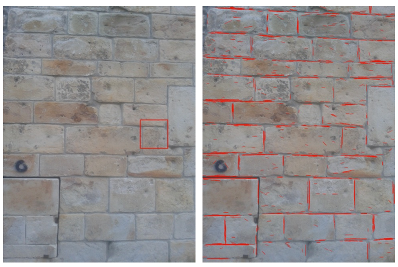
Our objective is to develop image-based tools that help automate the application of protocols for built heritage protection. In this paper, we address the image-based extraction of arrangement of the masonry for automatic classification. At first glance, one may think of applying automatic image-based granulometry techniques in order to delineate the individual stones and then infer the category from the geometric description of the delineated individual blocks. In the last two decades, image analysis for rock particles has become a hot topic of research, and a number of image systems have been developed for segmenting and measuring rock particles in different applications such as gravitational flows, conveyor belts, rockpiles, and laboratories, and some of them are under development [11]. Most of these systems try to segment the particles in the images in order to infer information about sizes and rely on the extraction of intensity gradient based edges.

In our case, the fully automatic, image-based delineation of individual stones can be very challenging mainly due to the characteristics of the objects to be delineated (delineation, in the context of our work, refers to the extraction of the outline of the construction blocks). These objects (built heritage) are exposed to the elements and, as a result, they suffer discolouring, cracking, erosion, and can even have vegetation growing in them. The image capturing is performed in an uncontrolled environment with lighting conditions possibly changing for different captures. Images can, therefore, have bright and dark areas, and shadows depending on the sun's position. In the case of bridge walls, the images can also contain reflections of the water. These undesired effects in the images generate intensity gradients that are, often, completely unrelated to the physical delineation of stone and, as a consequence, conventional edge-based methods are not useful since the signal-to-noise ratio is very low.

This motivates the development of a delineation method that does not rely on conventional edge detection methods. Our proposed framework for the semi-automatic delineation and classification of built heritage has two phases. In the first phase, the test image undergoes a series of processing steps in order to extract a set of straight line segments from which a statistical signature is inferred. This set of straight line segments constitutes a partial delineation of blocks. In the second phase the statistical signature is classified using the K Nearest Neighbor Classifier.

3 Proposed Delineation Framework

The aim of the proposed framework is to extract a set of straight line segments without relying on the intensity gradient edges. The first step of this framework is to convert the colour image into a grayscale image and partition it into regions of interest (ROI) of a predetermined size (approx. 300×300 pixels) for a 1280×960 image. Each ROI is processed independently. Figure 1.a shows the original image and an example ROI (marked in red). Figure 1.b shows the obtained delineation.

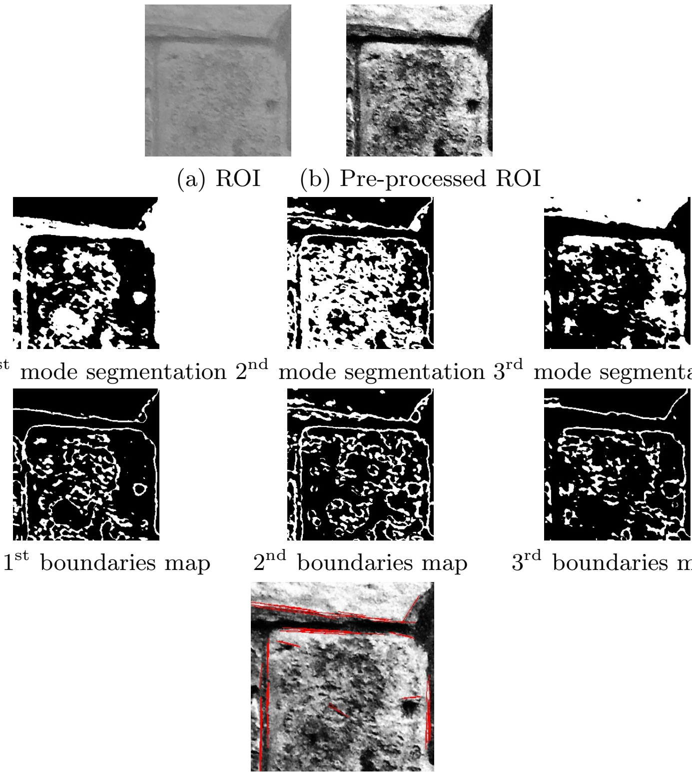


(a) Original **(b) Delineation**

Fig. 1. (a): Original image (scaled with example ROI marked in red). (b): The obtained delineation (a set of 2D segments marked in red) using the proposed framework.

3.1 Preprocessing of a ROI

Each grayscale ROI is prepared for processing by removing outliers and equalizing its histogram. Outliers are those intensities with a frequency below some threshold in the intensity histogram. Outlier pixels are smoothed out by inpainting, a method for removing damaged areas by taking the color and texture at the border of the damaged area and propagating and mixing it inside the damaged area [12]. Then, the histogram is equalized. One can notice that the images are not processed as a whole because, since these are images of objects that are outdoors, quite often it is not possible to control the lighting at the time of the capture and the image might contain regions with shadows, dark and bright areas, water reflections, and similar effects. Equalizing the histograms of the ROIs independently helps avoid these effects to some extent. Figure 2 shows the main steps related to the processing of a given ROI. See Fig. 2.a for our example grayscale ROI and Fig.2.b after removing outliers and equalizing its histogram.



(a) ROI (b) Pre-processed ROI

(c) 1st mode segmentation 2nd mode segmentation 3rd mode segmentation

(d) 1st boundaries map 2nd boundaries map 3rd boundaries map

(e) Extracted 2D segments associated with the three maps

Fig. 2. Main stages of processing a ROI. (a): An original ROI. (b): The preprocessed ROI. (c): Region segmentation using three modes. (d): Corresponding boundaries. (e): Extracted straight segments.

3.2 Processing of a ROI

The objective is to use straight segments to delineate the masonry. In our context, the straight 2D segments are not derived from the image gradients but from the boundaries of some detected regions in the preprocessed images. We have chosen to extract straight segments because this allows us to obtain additional information such as slope, length and position of the segment. The statistics of these geometric properties will be very valuable for subsequent classification tasks. The processing of a ROI is achieved following the next sequence of steps.

Region Segmentation Using Most Frequent Intensities. For each ROI, a number of its most frequent intensities are chosen. For each chosen intensity (mode), a binary image is created using an interval of intensities centred on the chosen intensity (mode) and with a certain radius. This binary image can be thought of as a bipartition of the original image. This bipartition will result in a region segmentation. The intuition behind using bipartition based on several modes is that we can usually find mortar or a dark shadow in between construction blocks and, thus, this image would go a long way toward delineating the blocks. In the case of a delineating shadow, this low intensity will quite often not be one of the most frequent intensities and, thus, it is included by default in the set of modes. In the case of binding mortar being present, although the construction blocks are often of different colors the mortar is originally of a uniform color. However, due to weathering and lighting effects, it will most likely take different intensities throughout the ROI. For these reasons, we use several intensity modes each performing a given bipartition, i.e., a region segmentation. In our work, we use three intensity modes (one always being 0, the two remaining modes are chosen from the histogram peaks) and a radius of 50 that was empirically selected. We stress the fact that the number of modes can be more than three. In our case, we find that three modes are enough in order to give accurate results.

By using the three modes and their support in grey-level scale (that may overlap), one can get three binary images. Figure 2.c shows the three binary images for the chosen three modes. These images correspond to three region segmentations. All of these three binary images are processed independently and their 2D line segments detected are all saved as part of the delineation set.

Extracting Boundaries by Removing Inner Patches. Once the image regions are segmented, it will be useful to extract their boundaries. To this end, we remove the inside of inner patches that may be present in the binary images so that the probabilistic Hough transform (PHT) [13] can segment the corresponding boundaries only. Inner patches are detected with an averaging filter. When this filter is run on a binary image, the result for each pixel is directly proportional to the number of white neighbouring pixels (belonging to the segmented region). Those pixels with a result above a certain threshold (depending on the size of the neighbourhood used) will be identified as being part of a patch and they will be removed. The usual method for extracting boundaries is to subtract the eroded binary image from the original. However, our experiments have shown that the method presented in this paper obtains better results for the purpose at hand. Figure 2.d shows the extracted boundaries obtained from the three binary images associated with the ROI.

Straight Segment Extraction. Once we get the binary boundaries, we apply a closing operation on them. The probabilistic Hough transform (PHT)[1] is then

[1] As an implementation of this algorithm, we use the CV_HOUGH_PROBABILISTIC method of the cvHoughLines2 function in OpenCV 2.3.1 [14].

used in order to extract the 2D straight line segments. The result of this process for the example ROI can be seen in Fig. 2.e. In this figure, we can observe that the line segments are close to each other in some parts of the delineation. This redundancy is due to the use of three segmented images that may have region boundaries close to each other.

3.3 Fusion of Results and Postprocessing

The final delineation and feature extraction is the set of all the segments detected in all three binary images and in all ROIs. The postprocessing of the delineation consists in joining collinear segments that are close to each other and removing small segments. Collinear segments are those for which the difference between their polar coordinates is below some predefined threshold. Two segments that have been classified as collinear are joined together if the distance between their closest end points is below some predefined threshold. Joining two segments together means that those two segments are replaced by a new straight segment delimited by the two furthest end points. Lastly, segments whose length is below the mean length have been removed from the set because their slope is not very reliable and they are often noisy. Figure 1.b illustrates the final delineation of the masonry (a set of 2D segments shown in red) associated with the image. We point out that this post-processing step is very useful for overcoming the fragmentation of the detected segments due to the ROI partition of the full image.

4 Performance Evaluation: Masonry Classification

In order to evaluate the usefulness of the presented delineation framework, we have exploited the delineation results in a classification task. One of the features the protocol for stone bridges evaluates is the arrangement of the masonry. The are three classes of masonry arrangement: the first class are blocks (usually irregular) not arranged in rows (Fig.3.(a)), the second are irregular blocks arranged in rows (Fig.3.(b)), and, third, regular (rectangular) blocks arranged in rows (Fig.3.(c)). We stress the fact that this classification problem is quite different from the scene categorization problem [15] where each class can refer to a different concept (sea, street, building, car, human).

Several statistics are extracted from the set of remaining delineation segments and are used as the predictor variables in the classifier. Regarding the lengths, these statistics are the maximum, mean, and standard deviation, where the lengths are expressed as a percentage of the image width. Regarding the slopes, the statistics used are the mean, standard deviation, and the percentage of slopes that are vertical (or very nearly vertical i.e. 90 degrees' slope). For slope differences (the slope difference for every pair of segments in the delineation is calculated), the mean, standard delineation, and the percentage of slope differences that are 0 or 90 degrees or very close to those marks. For the image of the processed delineation, the sums of the rows and the sums of the columns

Fig. 3. Examples of the three categories to be discriminated

are calculated and, for each, the maximum (expressed as the percentage of the width and the height, respectively), the mean, standard deviation, the absolute difference between the mode and the mean, and the maximum to mean sum ratio statistics are used. Altogether, every image is described by 19 features that summarize statistics about the extracted straight segments.

There are 86 instances (i.e. bridge wall images captured at different sites); 33 belong to the first class, 15 belong to the second, and the remaining 38 belong to the third class. After performing leave-one-out cross-validation, the Nearest Neighbour classifier has achieved an average accuracy percentage of 86.05%. This result is satisfactory but it also suggests that there is room for improvement. The proposed framework can be improved by reducing the noisy segments and improving the detection of the desired segments that are not being detected. The precision is 0.91, 0.62, and 0.91 and the recall is 0.96, 0.66, and 0.84 for the three classes, respectively. This means that the classifier's success rate can also be increased by improving the characterisation of the second class through the addition of features that contribute to discriminate this class from the others. The F1-score for each class is 0.94, 0.64, and 0.87, respectively.

5 Conclusions and Future Work

This paper has presented the foundational work carried out for the development of a built heritage analysis and classification ICT tool that will automate the application of the new protection protocols under development with the objective of speeding up the protection process. The new method introduced in this paper for the semi-automatic delineation of the stone masonry has been evaluated using a classification task. The obtained results are very satisfactory and show that the obtained delineation provides meaningful information.

Future work will focus on improving the delineation by removing as many noisy segments as possible; noisy, in this context, meaning all segments that do not belong to the delineation. The other research line will focus on the selection of the most relevant features for improving the classification.

References

1. Kersten, T.P., Lindstaedt, M.: Image-based low-cost systems for automatic 3D recording and modelling of archaeological finds and objects. In: Ioannides, M., Fritsch, D., Leissner, J., Davies, R., Remondino, F., Caffo, R. (eds.) EuroMed 2012. LNCS, vol. 7616, pp. 1–10. Springer, Heidelberg (2012)
2. Moussa, W., Abdel-Wahab, M., Fritsch, D.: Automatic fusion of digital images and laser scanner data for heritage preservation. In: Ioannides, M., Fritsch, D., Leissner, J., Davies, R., Remondino, F., Caffo, R. (eds.) EuroMed 2012. LNCS, vol. 7616, pp. 76–85. Springer, Heidelberg (2012)
3. Stanco, F.: Digital Imaging for Cultural Heritage Preservation: Analysis, Restoration, and Reconstruction of Ancient Artworks. CRC Press (2011)
4. Godin, G., Beraldin, J.A., Taylor, J., Cournoyer, L., Rioux, M., El-Hackim, S., Baribeau, R., Blais, F., Boulanger, P., Domey, J., Picard, M.: Active optical 3-D imaging for heritage applications. IEEE Computer Graphics and Applications (2002)
5. Muller, P., Wonka, P., Haegler, S., Ulmer, A., Gool, L.V.: Procedural modeling of buildings. ACM Transactions on Graphics 25(3), 614–623 (2006)
6. Pasko, G., Pasko, A., Vilbrandt, T., Filho, A., da Silva, J.: Digital interpretation of cultural heritage: 3D modeling and materialization of 2D artworks for future museums. International Journal of the Inclusive Museum 3(1), 63–80
7. Remondino, F.: Heritage recording and 3D modeling with photogrammetry and 3D scanning. Remote Sensing 3, 1104–1138 (2011)
8. Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J., Fulk, D.: The digital michelangelo project: 3d scanning of large statues. In: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 2000, pp. 131–144 (2000)
9. Ikeuchi, K., Oishi, T., Takamatsu, J., Sagawa, R., Nakazawa, A., Kurazume, R., Nishino, K., Kamakura, M., Okamoto, Y.: The great buddha project: Digitally archiving, restoring, and analyzing cultural heritage objects. International Journal of Computer Vision 75(1), 189–208 (2007)
10. Rodríguez, A., Valle, J., Martínez, J.: 3d line drawing from point clouds using chromatic stereo and shading. In: Proceedings of the 14th International Conference on Virtual Systems and Multimedia, pp. 77–84 (2008)
11. Wang, W.: Rock Particle Image Segmentation and Systems. In: Pattern Recognition Techniques, Technology and Applications. I-Tech, Vienna, Austria, pp. 197–226 (2008)
12. Telea, A.: An image inpainting technique based on the fast marching method. Journal of Graphics Tools 9, 25–36 (2004)
13. Kiryati, N., Eldar, Y., Bruckstein, A.: A probabilistic hough transform. Pattern Recognition 24(4), 303–316 (1991)
14. Bradski, G.: The OpenCV Library. Dr. Dobb's Journal of Software Tools (2000)
15. Zhang, W., Qin, Z., Wan, T.: Image scene categorization using multi-bag-of-features. In: International Conference on Machine Learning and Cybernetics (2011)